358
7.9.70

z

Psychological Bulletin

EDITED BY

LYLE H. LANIER, UNIVERSITY OF ILLINOIS

CONSULTING EDITORS

STEUART H. BRITT
McCann-Erickson, Inc., New York
DORWIN CARTWRIGHT
University of Michigan
FRANK A. GELDARD
University of Virginia
JAMES J. GIBSON
Cornell University
DAVID A. GRANT
University of Wisconsin
WILLIAM T. HERON
University of Minnesota
ERNEST R. HILGARD
Stanford University
WILLIAM A. HUNT
Northwestern University
JEAN WALKER MACFARLANE
University of California
DONALD G. MARQUIS
University of Michigan
JOHN T. METCALF
University of Vermont
JAMES G. MILLER
University of Chicago
NEAL E. MILLER
Yale University
HELEN PEAK
University of Michigan
ROBERT R. SEARS
Harvard University
ROBERT L. THORNDIKE
Teachers College, Columbia University

VOLUME 47, 1950

PUBLISHED BI-MONTHLY BY
THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
1515 MASSACHUSETTS AVE., N.W.,
WASHINGTON 5, D. C.

BUREAU OF EDUCL. & PSYL. RESEARCH

[illegible] Edn.
[illegible]
[illegible] 7. 9. 76
[illegible] No J - 358

CONTENTS OF VOLUME 47

ORIGINAL CONTRIBUTIONS, GENERAL REVIEWS, DISCUSSIONS

Leadership, membership and organization: Ralph M. Stogdill. 1

The indirect assessment of social attitudes: Donald T. Campbell 15

An experimental analogue of repression. I. Historical summary: Anchard Frederic Zeller........................ 39

On the circularity of the law of effect: Paul E. Meehl........ 52

The nature of the galvanic skin response: Robert A. McCleary. 97

Homogeneity of variance and the Latin square design: Allen L. Edwards................................ 118

A qualification in the use of analysis of variance: Wilse B. Webb and Vernon W. Lemmon........................ 130

A survey of audio-visual aids for teaching of clinical psychology: Elias Katz.................................. 137

A note on London's historical survey of psychology in the Soviet Union: Gregory Razran........................ 146

A reply to Razran's note: Ivan D. London.................. 150

What constitutes a good textbook in social psychology?: Dorothy G. Park...................................... 152

Partial reinforcement: a review and critique: William O. Jenkins and Julian C. Stanley, Jr.......................... 193

The factorial analysis of animal behavior: Joseph R. Royce.... 235

A note on the use of the method of successive differences in empirical curve fitting: W. R. Garner................. 260

Four studies in psychology and social status: Richard Centers. 263

Some Psychological problems of economics: Samuel P. Hayes, Jr. 289

The misuse of chi-square—a reply to Lewis and Burke: Charles C. Peters.................................... 331

Some comments on "The use and misuse of the chi-square test": Nicholas Pastore........................... 338

On "The use and misuse of the chi-square test"—the case of the 2×2 contingency table: Allen L. Edwards............ 341

Further discussion of the use and misuse of the chi-square test: Don Lewis and C. J. Burke...................... 347

Studies in the psychology of children's drawings: II 1928–1949: Florence L. Goodenough and Dale B. Harris....... 369

On the validity of inferences drawn from the fitting of Poisson and negative binomial distributions to observed accident data: J. S. Maritz......................... 434

Recent research on pain perception: Ward Edwards 449
Statistical techniques in present-day psychodiagnostics: Ivan Norman Mensh 475
Rorschach validation: Some methodological aspects: Leonard I. Schneider 493
Critique of Hornell Hart's "Social science and the atomic crisis": Ralph H. Gundlach 509
Rejoinder to Gundlach's critique: Hornell Hart 513

BOOK REVIEWS

Thorndike's Personnel selection: S. Rains Wallace, Jr. 76
Chase's The proper study of mankind: Henry J. Meyer 78
Smith's Psychological studies in twin differences: Anne Anastasi 80
Johnson's Essentials of psychology: Edward E. Anderson 81
v. Schiller's Aufgabe der Psychologie: Frank A. Pattie 83
Machover's Personality projection in the drawing of the human figure: Lillian Wald Kay 84
Mayo's Some notes on the psychology of Pierre Janet: James G. Miller 85
Symonds' Dynamic psychology: Franklin J. Shaw 86
Thorpe and Katz's The psychology of abnormal behavior: Wilbur S. Gregory 87
Frank's Projective methods, and Bell's Projective techniques: Arthur Weider 89
Wolff's Diagrams of the unconsious: Fred Y. Billingslea 91
Harris' Modern trends in psychological medicine: Livingston Welch 93
Krech and Crutchfield's Theory and problems of social psychology: Leo Postman 155
Britt's Social psychology of modern life: Ross Stagner 162
LaPiere and Farnsworth's Social psychology: Arthur Jenness 163
Pollak's Social adjustment in old age: John T. Metcalf 164
Schaefer-Simmern's The unfolding of artistic ability: Heinz Werner 165
Andrews' Methods of pschology: Conrad G. Mueller 166
Syngg and Combs' Individual behavior: C. M. Louttit 170
Appicciafuoco's Sommario di psicologia: Raymond J. Corsini 172
Zipf's Human behavior and the principle of least effort: Helen Peak 172
Ryan's Work and effort: H. M. Johnson 174
Buros' Third mental measurements yearbook: D. A. Worcester 177

Super's Appraising vocational fitness by means of psychological tests: MILTON HAHN 178
Mathewson's Guidance policy and practice: FRANK M. FLETCHER, JR. 180
Donahue, Coombs and Travers' Measurement of student adjustment and achievement: HAROLD SEASHORE 181
Hurlock's Adolescent development: ROBERT P. FISCHER 183
Escalona's An application of the level of aspiration experiment to the study of personality: RICHARD T. SOLLENBERGER 185
Lief's The commonsense psychiatry of Dr. Adolf Meyer: LELAND W. CRAFTS 185
Kanner's Child psychiatry: FRED MCKINNEY 187
Dunbar's Synopsis of psychosomatic diagnosis and treatment, and Alexander and French's Studies in psychosomatic medicine: BERNARD F. RIESS 189
Postman and Egan's Experimental psychology: FRANK W. FINGER 272
Dennis' Readings in general psychology: NORMAN L. MUNN 273
Valentine and Wickens' Experimental foundations of general psychology: STANFORD C. ERICKSEN 274
Murphy's Historical introduction to modern psychology: DONALD M. JOHNSON 275
Humphrey's Directed thinking and Hutchinson's How to think creatively: HAROLD GUETZKOW 277
Swenson, Anderson and Stacey's Learning theory in school situations: HOWARD H. KENDLER 279
Mead's Male and female: GEORGENE H. SEWARD 280
Lindesmith and Strauss' Social psychology: GLEN HEATHERS ... 282
Frank's Courts on trial: STEUART H. BRITT 284
Johnson's Statistical methods in research: HAROLD A. EDGERTON 285
Dashiell's Fundamentals of general psychology: HELEN PEAK ... 356
O'Kelly's Introduction to psychopathology: SEYMOUR B. SARASON 357
Salter's Conditioned reflex therapy: O. H. MOWRER 358
Deri's The Szondi test: HELEN SARGENT 360
Korner's Some aspects of hostility in young children: NANCY BAYLEY 361
Brouwer's Student personnel services in general education: E. G. WILLIAMSON 362
Vernon and Parry's Personnel selection in the British Forces: DOUGLAS H. FRYER 363

Mursell's Psychological testing: FREDERICK B. DAVIS.......... 365
Underwood's Experimental psychology: an introduction: JOHN L. FINAN.......... 444
Stagner's Psychology of personality: FILLMORE H. SANFORD.... 446
Sheldon's Varieties of delinquent youth: WILLIAM A. HUNT..... 448
Anderson's The psychology of development and personal adjustment: VINCENT NOWLIS.......... 517
Guthrie and Edwards' Psychology: a first course in human behavior: RICHARD P. YOUTZ.......... 517
v. Holzschuher's Praktische Psychologie: die Primitivperson im Menschen: HANS H. STRUPP.......... 519
Thorndike's Selected writings from a connectionist's psychology: JOHN B. WOLFE.......... 520
Wever's Theory of hearing: K. D. KRYTER.......... 522
Rubin's Experimenta psychologica: KARL F. MUENZINGER..... 523
Crow and Crow's Educational psychology: SIDNEY L. PRESSEY. 524
Rosenzweig's (with Kogan) Psychodiagnosis: HOWARD F. HUNT. 525
Steckle's Problems of human adjustment: DWIGHT W. MILES... 527
McKinney's The psychology of personal adjustment: H. MELTZER.......... 528
Almy's Children's experiences prior to first grade and success in beginning reading, Gray's On their own in reading, and Witty's Reading in modern education: FRANCES ORALIND TRIGGS.......... 529

MISCELLANEOUS

Books and Materials Received..........95, 286, 366, 531
Index of Subjects.......... 535
Index of Authors.......... 537

Vol. 47, No. 1 January, 1950

Psychological Bulletin

LEADERSHIP, MEMBERSHIP AND ORGANIZATION[1]

RALPH M. STOGDILL
The Ohio State University

The present paper is concerned with a point of view regarding the relation of leadership to group organization. It represents one attempt within the Ohio State Leadership Studies[2] staff to clarify and systematize certain aspects of the leadership problem. Such clarification appears to be necessary as a preliminary step toward the development of an integrated program of research on leadership problems in formal organizations.

The pioneering work of Lewin (**9**), Moreno (**10**), and their followers has resulted in marked progress in the development of methods for studying leadership as a phenomenon of *groups*. However, comparable progress remains to be made in the development of methods for the study of leadership as an aspect of *organization*. Several factors appear to have operated as barriers to the development of scientific theory and method in this area. One is the lack of an adequate definition of leadership. A second is the fact that in much of the literature on leadership, the terms "group" and "organization" are used interchangeably or are defined in exactly the same terms. A third derives from two opposed theoretical approaches represented, on the one hand, by those theories of organization in which the leader is conceived as a symbol of authority or as an embodiment of superior personal traits, and on the other

---

[1] A cooperative contribution of the U. S. Navy, Office of Naval Research, and the Ohio State University Research Foundation. The opinions presented are those of the author, and should not be regarded as having the endorsement of the Department of the Navy. The paper was presented in part before the Midwestern Psychological Association Chicago, Ill., April 29, 1949.

[2] The Ohio State Leadership Studies are designed as a ten-year program of research on leadership problems in military, business, industrial, educational and civilian governmental organizations. The staff includes C. L. Shartle, Director; Alvin E. Coons, Melvin Seeman and Ralph M. Stogdill, Associate Directors; John Hemphill, Research Associate; Donald T. Campbell, Research Consultant; Richard T. Morris and Charles M. Westie, Research Assistants.

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated........................
Accs. No........................

hand, by a type of group-oriented theory in which leadership appears to be regarded as a manifestation of social pathology. A fourth, and related obstacle, results from a reaction of social scientists against the authoritarian principles advanced in many discussions of organization. Some social theorists appear to reject all concepts of organization as authoritarian; and some researchers appear reluctant to deal experimentally with such concepts as responsibility, authority, stratification and similar phenomena related to organization. It is beyond the scope or purpose of this paper to portray the magnitude of the latter two difficulties. Nevertheless, it seems relevant to recognize the fact that they are present and act to the detriment of scientific work in the field.

The Ohio State Leadership Studies are being conducted on the basis of these assumptions: (1) that group organization is a recognizable social phenomenon in our culture; (2) that as such it is a legitimate subject for scientific study; and (3) that the variables of organization can be isolated and defined so as to permit their scientific study. It is the purpose of the present paper to examine various concepts relevant to leadership and organization, and to develop a formulation of the problem which will suggest hypotheses that can be subjected to experimental test.

GROUPS AND ORGANIZATIONS

Wilson (**17**) has reviewed the important sociological literature relating to concepts of the social group. He reports that in "current sociological literature one finds no consensus as to the meaning of the *group*," and concludes that much experimental work is yet to be done in order to delimit the group concept in any satisfactory manner. An important step in this direction has been made by Hemphill (**5**), who has devised scales for the measurement of such group dimensions as size, permeability, stability, viscidity homogeneity of membership, and the like.

The most satisfactory definition available at the present time appears to be that of Smith (**15**) who defines a *social group* as "a unit consisting of a plural number of organisms (agents) who have collective perception of their unity and who have the ability to act/or are acting in a unitary manner toward the environment." Krech and Crutchfield (**8**) present a similar view. They state that "the criteria for establishing whether or not a given set of individuals constitutes a psychological group are mainly two: (1) all the members must exist as a group in the psychological field of each individual, i.e. be perceived and reacted to as a group; (2) the various members must be in dynamic interaction with one another."

A special kind of group is the *organization*. An organization may be defined as a social group in which the members are differentiated as to their responsibilities for the task of achieving a common goal.

Znaniecki (**18**) has reviewed the sociological literature relating to various concepts of organization. He stresses the fact that the terms *group* and *organization* are rather tenuous concepts, in that it is often difficult to determine whether a particular aggregate of persons constitutes a group, and that it may also be difficult at times to determine whether a particular group can be regarded as an organization. He points out that social organization

> . . . . can be realized only in a lasting "social group" or "association." Individuals belonging to such a group are aware that they will be regularly expected to perform certain actions, and some of them act as organizers, leaders, coordinators of the regular activities of others with reference to the common purpose. Not all of these individuals need be continuously active; indeed, in many groups a considerable proportion remain passive, acting only in reaction to the actions of others. The common purpose of the organized actions may be simple or complex.

Some of the consequences of distinguishing between the terms "group" and "organization" are the following: First, there is nothing in the term "group" which gives any clue as to the nature of leadership. Second, there is nothing in the group definition which provides any foundation for integrating leadership with group phenomena, except at a superficial level of social perception or interaction. Third, the group orientation can suggest research methods relating to leadership only in so far as the social group is defined in terms of organization. The concept of organization, however, with its implications for the differentiation of responsibility roles, does permit the study of leadership as an aspect of the relationships between members who are coordinating their efforts for the achievement of common goals.

A group may or may not have leaders. If it does have leaders, it is an organization, for at least some of the members are thereby differentiated from the others as to responsibility, or role expectation in relation to some common purpose. The members of a group may or may not have mutual responsibilities for a common task. If the members do have differentiated responsibilities in relation to common goals then the group is an organization—a particular kind of group. The continued presence of leaders and of responsibility differentiations in relation to group goals are indicative of organization. It may not always be easy to determine the exact point at which a group emerges into an organization.

LEADERSHIP AS AN ASPECT OF ORGANIZATION

The following definition of leadership may serve as a starting point for discussion. Leadership may be considered as the process (act) of influencing the activities of an organized group in its efforts toward goal setting and goal achievement. The definition of leadership relates it directly to the organized group and its goal. It would appear that the minimal social conditions which permit the existence of leadership are the following:

1. A group (of two or more persons).
2. A common task (or goal oriented activities).
3. Differentiation of responsibility (some of the members have different duties).

There are innumerable other group and situational factors which may influence leadership in varying degrees, but these appear to be the minimal conditions which will permit the emergence of leadership. There must be a group with a common task or objective, and at least one member must have responsibilities which differ from those of the other members. If all members perform exactly the same duties in exactly the same way there is no leadership. A leader then is a person who becomes differentiated from other members in terms of the influence he exerts upon the goal setting and goal achievement activities of the organization.

The foregoing discussion suggests that leadership cannot emerge unless the members of a group assume different responsibilities. It has been suggested that group organization also is founded upon differentiation of responsibility. It would then appear that leadership and organization are derived from a common factor or, viewed from a different light, that leadership is an aspect of group organization. This view has been expressed in various forms by writers in the field of business organization. Davis (3), for example, states that the

> . . . . development of organization structure is largely a problem in the division of responsibility, and involves two important problems: (1) the grouping of similar functions to form the various organization elements in a manner that will promote effective cooperation, and (2) the determination of the proper relationships between functional groups and organization elements, with a view to promoting both cooperation and effective executive leadership.

The definition of leadership does not specify how many leaders an organization shall have, nor whether the leadership influence of an individual is continuous or intermittent, nor whether the influence of the leader shall be for the welfare or detriment of the organization and its members. It merely specifies that leaders may be differentiated from other members in terms of the extent to which they influence the activities of the organization in its efforts toward the achievement of goals. The definition of effective and ineffective leadership is an additional problem.

ASPECTS OF RESPONSIBILITY

Brown (2) in a challenging analysis of organization maintains that "An enterprise is a mosaic of its individual responsibilities. The sum of them must exactly equal the whole requirement of administration." He continues, "Responsibility is that part in administration which is assigned to a particular member of an enterprise. Its definition is an act of organization."

Responsibility cannot be regarded as a simple or uncomplicated variable. Jucius (7) writes,

> By responsibility is meant, first, the obligation to do an assigned task, and, second, the obligation to someone for the assignment. But what is meant by obligation and how far does it extend? This implies a willingness to accept, for whatever rewards one may see in the situation, the burden of a given task and the risks which attend in the event of failure. Because of the rewards and penalties involved, it is highly essential to specify the limits of responsibility.

Formal organization can seldom define all the possible variations of responsibility and personal interaction to be expected of all members in all situations. Nevertheless, organization appears to be founded upon a basic system of stable expectations regarding differential responsibilities and relationships among the members. This is not a one-way process. That is, it is not the organization alone which sets up role expectations for its members. The members set up expectations for each other and for the organization as a whole. It is assumed for purposes of the present discussion that this principle applies not only to stratified organizations, such as military and industrial establishments. It applies as well to membership in any organized group, whether it be a business, political, educational, religious, fraternal, or social organization and regardless of size, stratification, purpose, or member characteristics. The essential relationship which makes possible the conduct of organized group activities is a differentiation of responsibility roles among the members. Without this there is no possibility of coordination or of leadership toward goal achievement. The very process of organization defines the responsibilities of the members and thereby the formal leadership of the group. It is true that in some organized groups, such as recreational groups, the responsibilities of members may appear to be vaguely defined. However, this is not equivalent to saying that no responsibilities exist.

Responsibility, in its broadest scope, defines not only the duties for which a member is accountable; it defines also the persons to whom and for whom he is accountable in the discharge of his duties. In doing so, it also defines a member's formal status, or location in the organization hierarchy. Authority and formal status systems in organization are but aspects of the division of responsibility.

Responsibilities in a systematic organization are determined by the assignment of persons to particular positions, the duties of which are outlined in an organization manual or organization chart. In less systematic organizations the responsibilities of a particular job or position may be determined by on-the-spot instructions, by general hints or by unverbalized assumptions. In a systematic organization an individual's *work patterns* (the tasks he *actually* performs) will correspond fairly closely with his *responsibility patterns* (the tasks he *is supposed* to perform). However, as the mission and activities of the organization change there will be found in many instances an increasing discrepancy between the tasks being performed and the responsibilities originally outlined and defined.

ATTRIBUTES OF ORGANIZATION

The studies of Roethlisberger (12) and others have directed attention to the factor of informal groups within formal organization. Informal organization, as usually defined, refers to the friendship groups and cliques—based upon close association, mutual interests or antagonisms, and the like—which develop within formal organization. It has been pointed out by Homans (6) that this conception is too narrow, since what is informal in a factory may be formal in a primitive society. Firey (4), who defines informal organization in terms of schism, presents a more useful approach to the problem. He maintains that "if we regard behavioral conformity, in which interactional processes are highly repetitive and synchronized, as the overt counterpart of a social system, then behavioral nonconformity may be taken as the overt counterpart of schism within a system."

An organization in operation seldom corresponds exactly with the organization model as charted. The intervention of human social factors and other influences result in the emergence of informal organization, that is, in the development of work patterns and interaction patterns which do not correspond with responsibility patterns.

It would appear then that there are two fundamental sets of variables which define the operations of an organized group. These are:

1. *Variables which define formal organization.* These are:
   a. Responsibility variables (the work one is expected to do).

   - b. Formal interaction variables (the persons with whom one is expected to work).
2. *Variables which define informal organization.* These are:
   - a. Work performance variables (the tasks one actually performs).
   - b. Informal interaction variables (the persons with whom one actually works).

If we regard the variables listed above as basic variables of organization, we can also regard them as variables of membership and of leadership. In other words, an organization can be studied in terms of these four types of variables: responsibilities, work performances, formal interactions and informal interactions. Leadership can also be studied in terms of the same variables.

Responsibility variables define the duties that the members are expected to perform. The responsibilities of a given position may remain the same, whether A or B occupies the position. *Work performance* variables are defined by the tasks performed and by the methods of their performance. Individual A may accept a position previously occupied by B. The responsibilities as defined by organization charts and manuals may remain the same, but the tasks actually performed by A may differ somewhat from those performed by B, and the methods of performance may vary markedly.

Formal interaction variables define the persons to whom and for whom the members are accountable, as well as others with whom they are expected to cooperate, in the discharge of their responsibilities. *Informal interaction* variables are defined by the persons with whom the members actually work and cooperate in the performance of their tasks.

Informal organization comes about as a result of the development of discrepancies (a) between work performance and responsibilities as defined and (b) between informal interactions and formally defined interactions. Thus leadership is ever confronted with the task of reconciling discrepancies—discrepancies between what ought to be done and what is being done, between goals and achievements, between organizational needs and available resources, between the needs of individual members and the requirements of organization, between formal lines of cooperation and informal patterns of cooperation.

An organization in action comprises a complex of many variables in interaction. In making a pictorial representation of a business organization, the usual procedure is to plot the division of formal responsibility on a two dimensional chart. The horizontal dimension of the chart shows the division of responsibility for various kinds of work. The vertical dimension of the chart shows the division of responsibility for different levels of decision-making, and indicates the persons to whom one is accountable and those for whose performance one is accountable in the discharge of duties. This dimension defines the formal authority and status systems of the organization. Level (position in the organization hierarchy) and function (kind of work performed) are not independent dimensions. Although functions tend to differ from level to level, there is considerable overlap. Results from the Ohio State Leadership Studies (**13**, **16**) have shown that the functions of top leadership tend to be supported at each lower level in the leadership structure by increasingly more detailed and routine work in the same functions.

Personal interaction can also be conceived as varying in both horizontal and vertical directions. The horizontal dimension is defined by the range (number) of members with whom an individual interacts. Some persons tend to work alone or with single individuals, while others are observed to work with large numbers of persons. The vertical aspect of personal interaction is defined by the number of strata (echelons) above and below his own in which a member works with others. Some persons may be observed to work only with others at the same level in the organization. Others tend to work only with subordinates, and still others tend to work only with superiors. These tendencies may or may not represent expression of individual differences in social interaction patterns. Results obtained thus far in the Ohio State Leadership Studies suggest that these patterns of interaction may be determined in part by the functions served by various types of positions. Technical consultants and staff aides tend to spend more time with superiors. Members in supervisory positions are observed, as would be expected, to spend more of their time with assistants, and subordinates. Members in coordinative positions tend to spend time with superiors and subordinates, as well as with associates at the same level in the organization. A member's function or duties may determine to a considerable degree which persons in the organization he may influence, as well as the nature of the influence that he can exert.

GROUP ORGANIZATION DEFINES AND DELIMITS LEADERSHIP

The very process of defining responsibility serves to structure and delimit the role that the leader may play in the organization. He cannot perform all the duties of all the members. His own accomplishment is therefore dependent upon the performance of others. His responsibilities are circumscribed by the outlined procedures and delegated responsibilities necessary for the achievement of stated goals.

Each member must work within the organizational framework which defines the limits of his participation (how far he ought to go and beyond which he ought not to go) in performance of duties. It also sets the requirements for his cooperation with others and defines his relationships with his superiors and subordinates. This organizational structuring is not viewed alike by all persons. To some it appears as a barrier to participation or recognition. To others it appears as a prod and stimulus to greater effort and participation. For still others it provides a secure and comfortable sphere of activities and working relationships. Organization, therefore, in defining the responsibilities and working relationships of its members, sets up barriers to participation, as well as facilitating it.

Even as the organization sets boundaries by providing a framework within which members discharge their responsibilities, so the individual presents various barriers to the influence of the organization upon his own behavior and reactions. Some members may be limited in capacity to discharge their responsibilities, while others who are highly skilled in the techniques of their responsibilities are limited in capacity to interact with others. Each member carries into the organization his past experiences, his needs, ideals, personal goals, and committments to other organizations, which may modify and determine his capacity for participation. It would appear that the extent to which the behavior of different members is determined by the characteristics of the group represents a continuum from little to great, and also, that the extent to which the behavior of the different individuals determines the behavior of groups may be conceived as representing a similar continuum.

It becomes apparent that a study in leadership represents a study of relationships, of variables in interaction. According to Pigors (**11**), a study of leadership must consider: (1) the leader, (2) the members as individuals, (3) the group as a functioning organization, and (4) the situation.

All organizations operate within a larger cultural and environmental framework. No organization can escape entirely the influence of the external situation. The organization may be influenced by the availability of resources, by changes in the social order of which it is a part, by competition of other organizations for the participation, resources or loyalty of its members, and by innumerable other factors outside the control of the organization itself. These factors also influence the leadership of the group.

LEADERSHIP AND EFFECTIVENESS OF ORGANIZATION

According to Barnard (**1**) the persistence of cooperation depends upon two conditions: (a) effectiveness, the accomplishment of cooperative purpose, and (b) efficiency, the satisfaction of individual motives. Thus, although in many situations it may appear desirable to effect a maximum of goal achievement with a minimum of organizational expenditure, such a procedure might jeopardize the welfare or morale of the members. It then becomes evident that there are many situations in which organization is confronted by a complex of contradictory factors which must be considered in arriving at a decision. It also becomes apparent that the effectiveness of an organization cannot always be evaluated in terms of the degree to which it has attained its objectives. It may be necessary first to evaluate the goals and objectives themselves or the cost of their attainment. A carefully thought out discussion of factors to be considered in setting organizational goals, arriving at decisions, and evaluating the success of an organization has been presented by Simon (14). He states "The accomplishment by an administrative program of its organizational goals can be measured in terms of *adequacy* (the degree to which its goals have been reached) or of *efficiency* (the degree to which the goals have been reached relative to the available resources)." Simon, in agreement with Barnard, maintains that the criterion of adequacy alone is not valid as a measure of group accomplishment. He observes that "the fundamental criterion of administrative decision must be a criterion of *efficiency* rather than a criterion of *adequacy*. The task of administration is to maximize social values relative to limited resources."

If organizational goals are employed as reference points in evaluating effectiveness, then the goals themselves must be subject to evaluation. In addition, the cost (human or material) of goal attainment must be considered as a factor in evaluation. Both Barnard and Simon imply that organization cannot be regarded as a unit in isolation—or as a law unto itself. The motive of organization is the creation of social value or goods for its members, and these values bear some significant relation to the values of society in general.

Since leadership is related to the determination of group goals, it becomes apparent that the leader is seldom a free agent. In influencing the activities of the organization in its striving toward goal achievement he must consider certain social values, not only in relation to the members, but in relation to society as well. If he ignores the welfare of the members he is likely to lose their following. If he ignores the welfare of society he is likely to lead his group into difficulty. Thus leadership is subject to determination by factors which are external to the organization, as well as by internal group factors.

THE DEFINITION OF LEADERSHIP

The definition of leadership as a process of influencing the activities of an organized group in its task of goal setting and goal achievement should perhaps be reexamined. Does it define leadership? What are its implications? Admittedly, it defines only at a high level of generality. Certainly it does not include all social acts and influences, but it is nevertheless, an inclusive rather than a restrictive definition of leadership. Even so, it is more restrictive than most of those attempted in the recent literature. The definition restricts leadership to influence within the organized group. It does not imply domination or direction of others, nor any technique of influence; nor does it specify any particular member who should be regarded as a leader. The definition permits the study of any member of an organization to determine the extent of his leadership influence, and permits consideration of the possibility that every member may contribute toward determining the leadership of the organization.

The definition carries the implication that leadership may be not so much an attribute of individuals as an aspect of organization. In this respect it parallels the concept of authority. It is generally recognized that an executive in a business concern has authority in relation to his employees only during the time they are working as members of the organization. His authority does not extend outward into the direction of their personal or social lives. Nor does his position as an executive give him authority over other persons who are not members of his organization. In other words, authority is a relationship that exists between persons who are mutually participating as members of an organized enterprise. Authority is not an attribute of one or a few persons. Authority is an interactional process by means of which the organization defines for each individual the scope for action he has in making decisions, carrying out responsibilities, and enlisting the cooperation of others. The authority of any single individual will be largely circumscribed and defined by the authority of others, and at the same time, his own degree of authority will in part determine the authority of others.

Leadership appears also to be determined by a system of interrelationships. As such it must be regarded as an aspect of organization, just as authority is a derivative of organization. If leadership is determined by a system of interacting variables, then each of the several dimensions of responsibility and personal interaction might be conceived as representing a gradient of influence. If so, then it should be possible to measure leadership influences in terms of these dimensions.

Some members may be regarded as rating higher than others in leadership by virtue of the fact that they have responsibility for making decisions which exert a marked influence upon the activities of the organization. Some members may influence the activities of the organization as a result of personal interaction with other members, even though they do not hold positions of high level responsibility. Some members may rate high in both types of influence. It would not be expected that any organization could be found in which all influence is exerted by a single member. It would rather be expected that all the members of the organization could be ordered or ranked to some degree in terms of the influence they exert in various dimensions. The proposal to measure leadership in terms of the influence exerted by individuals may appear to contradict the statement that leadership is an aspect of organization rather than an attribute of individuals. But this is not a necessary conclusion. It was pointed out that authority is generally understood to be an aspect of organization. However, it can be observed that some members exercise more authority than others. The judgment can also be made that some persons have "too much" or "too little" authority. Such observations indicate an evaluation of conditions relative to various factors in the organization. In the same way it can be observed that member A exerts more leadership influence in some situations; while members B, C, and D exert more influence in determining activities of the organization in other instances. It may be that the leadership of A is circumscribed by the leadership of B, C, and D who are in competition with him; or it may be that the leadership of A is dependent upon the supporting leadership of B, C and D. In either event, the leadership influence of any one member is determined in part by the leadership exerted by others, and the balance may change from time to time.

SUMMARY

An organization is composed of individuals. Its existence is dependent upon the cooperation and performance of individuals who play different roles. Measures of authority, leadership, and the like, are but measures of aspects of organization, even though the measurements are made in terms of members and the relationships among members. Leadership exists only in so far as individuals, as members of organization, are differentiated as to the influence they exert upon the organization; and the leadership influence of any one member will be determined to a large degree by the total leadership structure of the organization. It is for this reason that leadership has been here defined in terms of influence upon the activities of the organization, rather than in terms of influence upon persons.

The advantages of this formulation of the leadership problem are as follows: First, it removes leadership from the broad, vaguely defined realm of social interaction in general, and integrates it with the basic variables which describe an organized group. Second, and more important, is the fact that it suggests the development of methods for studying leadership as an aspect of work performance, work methods and working relationships.

An attempt is being made to develop such methods for the Ohio State Leadership Studies. For example, the goals and structure of organization and the responsibility patterns of members are determined by examining organization charts and manuals and by interviews with members of the organization. Work patterns are determined by modified job analysis procedures. Sociometric methods are employed to determine working relationships between the members and to chart the informal organization. The social values and role concepts of leaders and members are studied by means of attitude scales. These methods are supplemented by various check lists and rating scales.

In conclusion, a word of caution may be in order. The present paper has been concerned with a search for the minimal factors which will permit a functional integration of the concepts: leader, member and organization. In attempting to isolate these minimal common elements, many other important factors associated with leadership and group functioning have been excluded as not contributing to this central purpose. The present formulation represents merely one segment of a set of hypotheses to be subjected to experimental test.

BIBLIOGRAPHY

1. Barnard, Chester I. *The functions of the executive.* Cambridge: Harvard Univ. Press, 1938.
2. Brown, Alvin. *Organization of industry.* New York: Prentice-Hall, 1947.
3. Davis, Ralph C. *Industrial organization and management.* New York: Harper, 1940.
4. Firey, Walter. Informal organization and the theory of schism. *Amer. sociol. Rev.*, 1948, **13**, 15–24.
5. Hemphill, John K. *Situational factors in leadership.* Ohio State Univ., Bur. Educ. Res. Monogr. 31, 1949.
6. Homans, G. C. A conceptual scheme for the study of social organization. *Amer. sociol. Rev.*, 1947, **12**, 13–26.
7. Jucius, Michael J. *Personnel management.* New York: Irwin, 1947.
8. Krech, David, & Crutchfield, Richard S. *Theory and problems of social psychology.* New York: McGraw-Hill, 1948.
9. Lewin, Kurt, Lippitt, Ronald, & Escalona, Sybille K. *Studies in topological and vector psychology I.* Univ. Iowa Stud. Child Welf., **16**(3), 1940.
10. Moreno, J. L. *Who shall survive?*

Washington: Nervous and Mental Diseases Publ. Co., 1934.

11. Pigors, Paul. *Leadership or domination.* New York: Houghton Mifflin, 1935.
12. Roethlisberger, F. J., & Dickson, William J. *Management and the worker.* Cambridge: Harvard Univ. Press, 1939.
13. Shartle, Carroll L. Leadership and executive performance. *Personnel,* 1949, **25**, 370–380.
14. Simon, Herbert A. *Administrative behavior.* New York: Macmillan, 1947.
15. Smith, Mapheus. Social situation, social behavior, social group. *Psychol. Rev.*, 1945, **52**, 224–229.
16. Stogdill, Ralph M., & Shartle, Carroll L. Methods for determining patterns of leadership behavior in relation to organization structure and objectives. *J. appl. Psychol.*, 1948, **32**, 286–291.
17. Wilson, Logan. Sociography of groups. In G. Gurvitch and W. E. Moore (Eds.), *Twentieth Century Sociology.* New York: Philosophical Library, 1945. Pp. 139–171.
18. Znaniecki, Florian. Social organization and institutions. In G. Gurvitch and W. E. Moore (Eds.), *Twentieth Century Sociology.* New York: Philosophical Library, 1945. Pp. 172–217.

Received July 31, 1949.

THE INDIRECT ASSESSMENT OF SOCIAL ATTITUDES

DONALD T. CAMPBELL
The Ohio State University

In the problem of assessing social attitudes, there is a very real need for instruments which do not destroy the natural form of the attitude in the process of describing it. There are also situations in which one would like to assess "prejudice" without making respondents self-conscious or aware of the intent of the study. At the present time there are few if any indirect tests which could confidently be used for either of these purposes. There are, none the less, a considerable number of techniques that have been partially explored and validated. It is the purpose of this paper to survey such techniques and to present a point of view with regard to the problem of indirect measurement.

Current interest on the part of social psychologists in the indirect assessment of attitudes is perhaps primarily an aspect of the larger projective test movement in personality study. However, as will be seen in the course of this survey, there has been an interest in this approach from the very first efforts in attitude measurement (**81**), anticipating by some ten years the current interest in "projective techniques."

The terms "indirect" and "projective" have been used to refer to both *disguised*, and to *non-structured* measurement efforts. Using these two terms alone, one could distinguish four types of tests:

1. *Non-disguised-structured:* the classic direct attitude tests of Thurstone (**78**), Likert (**49**), *et al.*
2. *Non-disguised-non-structured:* the free-response interview and questionnaire approaches, the biographical and essay studies.
3. *Disguised-non-structured:* the typical "projective" techniques.
4. *Disguised-structured:* tests which approximate the objective testing of attitudes.

It is with the latter two categories that the present review is concerned, although some items that properly belong in category 2 will be included where they represent deliberate efforts at "projective" attitude testing.

While a formal division of content on these two criteria suffices, the writer will, in discussion, use a third which overlaps but does not duplicate the other two. This third criterion is that of dependence upon voluntary self-description as opposed to diagnosis based upon differential performance in an objective task. Upon this latter criterion rather than "structuredness" will rest the primary distinction between our two types of methods.

DISGUISED-NON-STRUCTURED TESTS OF SOCIAL ATTITUDES

In this category are included those techniques which offer the respondent opportunity for the spontaneous expression of attitudes in an ambiguous or non-structured setting. Most of these techniques are borrowed quite directly from well established clinical tools, and will be classified on that basis.

Approaches based on the Thematic Apperception Test. Though by no means the first, the indirect test most widely cited in previous surveys of the literature (**18**, **42**, **54**, **82**) is that of Proshansky (**65**). He intermingled ambiguous pictures of labor situations with the more usual T.A.T. scenes. The pictures were presented to a group by means of slides, with each person being asked to write for two and one-half minutes on what he thought the slide represented. Slides were shown for only five seconds. Proshansky found that ratings made from the resulting descriptions correlated .77 and .67 with a direct verbal scale of attitudes toward labor. A more elaborate proposal for the use of a similar technique with attitudes toward Negroes and Jews has been made (**51**), but as yet, no results are available.

Specially designed Thematic Apperception pictures have also been used by Frenkel-Brunswik, Sanford and Levinson in their extensive study of the personality correlates of prejudice (**24**). In this research the purpose was not so much to measure prejudice as to get a more detailed and qualitative picture of its expression. Nevertheless, the complicated interrelationships they found qualify the use of such pictures as attitude measuring instruments. For example, many prejudiced women told warmer and more sympathetic stories to a picture of a Negro mammy than they did to pictures of an elderly white woman (**23**). While such a finding is consistent with personality theory, it points to the danger of an over-simplified one-to-one interpretation of such material.

Essentially identical in technique is the "Human Relations Test" developed by the staff of the College Study in Intergroup Relations (**15**). This consists of ten drawings, each portraying an ambiguous intergroup contact. For example, a crowded street-car scene and a basketball scene are used. The respondents are asked to write a short story about each scene, telling what happened, how it came about, and what will happen next. The test is available on printed cards or projection slides, and since it is being used by a number of cooperating colleges, and at a wide variety of age levels, there is some promise of a standardized test emerging.

Johnson (**39**) has successfully studied the development of Anglo-Spanish attitudes in the Southwest through stories told by children to a specially designed series of pictures. He used six carefully selected conflict situations. These were duplicated in three forms: with all Anglo characters (for use with English-American children), with all Spanish characters (for use with Spanish children), and with mixed Anglo-Spanish characters (for use with both groups). Attitudes were assessed by contrasting responses to identical situations when they involved Anglo-Spanish conflict and when depicted by their own group members only. Quantification was achieved by having two judges categorize individual responses on a number of dimensions. Reliability coefficients for six sub-groups were over .90.

Approaches utilizing doll play techniques. In the first explicit attempt to use a projective technique in the assessment of social attitudes, Dubin (19) utilized toys in a fashion similar to the "play techniques" of the clinical psychologist. Using ten adult respondents, he asked them to "construct on this table a dramatic scene or scenes of the world as you see it today" and later "Now make a dramatic scene or scenes of the world as you would like it to be." From report and pictures of these scenes, three judges were able to estimate answers on 21 direct attitude questions dealing with labor, the Negro, internationalism, etc. with an average rank order correlation of .49.

For use with children, Arnold Meier has developed an interviewing aid, the "What Would You Do?" test, which is "projective" in the present sense. Doll cut-outs including minority children are manipulated on background drawings which depict such situations as entering the home, school scenes, etc. (15).

Similar to these approaches is the "movie story game" developed by Evans and Chein (21) for use with children from eight to twelve. On a miniature stage, Negro and white dolls are manipulated, with the child being asked at various points what the identified doll would say. Particular attention is paid to the patterning of segregation responses with acquaintances and with strangers. Preliminary studies indicate effective disguise and general meaningfulness for the test.

Hartley and Schwartz (33) have combined doll-play with pictorial material in the investigation of intergroup attitudes of children five to seven years of age. Montage background compositions carried characteristic symbols of Jewish religion in one case, Catholic in another, and in a third indicated typical middle class surroundings with no indication of religion. Identical family sets of dolls were placed upon these three backgrounds, and all were available for use by the child in playing out situations, such as a birthday party, school bus, meals, etc. In the preliminary report it is indicated that children identify with some accuracy the religious symbols, and that their play indicates intergroup attitudes in a meaningful way.

Sentence completion tests. At the Ohio State University, Shirley Wilcox Brown is using a modification of Rotter's test (68) in investigating attitudes toward the Negro. A 40-item schedule has been prepared, in which are imbedded some 20 sentence fragments dealing specifically with the Negro problem. Examples of relevant and neutral items are as follows:

1. I feel . . .
2. Skin color . . .
3. I hate . . .
4. Maybe . . .
5. Some lynchings . . .
6. The K.K.K. . . .
7. It seems to me that segregation . . .
15. Negro body odor . . .
37. Racial intermarriage . . .

In this test, no real effort at dissembling is made. The cooperating respondent is asked to volunteer a description of his own attitudes, with no more disguise than in the direct tests. The instructions read "Complete these sentences to express *your real feelings*. Try to do every one. Be sure to make a complete sentence!" Preliminary applications of the test indicate its power to discriminate between criteria groups. Scoring has been done by both coding and rating procedures. The indications are that the test can provide a reliable measure of attitudes.

Another twist to the sentence completion notion is being used in the study of personality and prejudice among school children conducted by Else Frenkel-Brunswik and Harold E. Jones (22). Here the test avoids all mention of minority groups, but provides stereotypic statements which may be completed with names of various minority groups or others. Examples of items are:

Are there some people who are mean? WHAT PEOPLE?
It would be better if more of a certain type of people were allowed to come into the United States. WHAT PEOPLE?
Some people are poor and it is their own fault. WHAT PEOPLE?

This test has elicited mention of foreign and minority groups from about one third of the children to which it has been administered. From another third or so come one or more anti-prejudice statements. A portion of the students make no responses classifiable in either way, and are thus not effectively evaluated. Using a net score (subtracting anti-prejudice responses from the total of prejudiced ones) corrected reliability figures run around .6 to .8 and correlations with a highly reliable direct test are on the order of .5. The approach is most satisfactory for the comparison of groups of respondents, and for the evaluation of the relative salience and extremity of attitudes toward different minority groups. In addition, unique data on the uniformity of stereotyping are provided.

While these two sentence completion tests make a deliberate effort to achieve a projective or indirect character, they are not too different from any free-response questionnaire on attitudes. Compare, for example, Zeligs' approach in which school children were asked to "write the most interesting true sentence" they knew about each group, within a one minute limit (84).

Miscellaneous non-structured techniques. The Rosenzweig Picture Frustration Test has also been modified for the measurement of social attitudes, a task to which it is obviously appropriate. This test requires the respondent to fill in the balloons in a series of cartoon drawings involving frustrating face-to-face inter-group contacts (6).

Probably belonging within the limits of our general category is the study by Fromme (25). He presented to the respondents five political cartoons, each with four alternative captions, covering a wide range of pro and con opinion. The respondent was asked to pick the best caption, and this choice, plus the discussion resulting, was utilized in a qualitative analysis of attitude structure.

The procedures just discussed are all indirect and "projective" in current usage of the terms. At the risk of some oversimplification, the assets they have in common may be emphasized and the stage set for contrasting them with the second group of procedures.

While tests such as these may be in part disguised, in that the experimenter does not tell the respondent his real purpose and may indeed substitute a false justification, their primary asset is that of securing an expression of attitudes in a more natural and spontaneous form, allowing opportunity for each individual to "project" upon a neutral screen his own integration of the problem. In contrast, the usual direct paper-and-pencil attitude test, requiring the endorsement of prepared items, may be said to force artificially the expression of attitudes into a preconceived and common mold. The great advantage of these non-structured tests is one of freedom, rather than of disguise. While some respondents may complete their assignments unaware of the experimenter's interest, in a tense situation one could hardly use any of the above devices expecting to get unconscious or uncensored expressions of attitude from uncooperative respondents. Essentally these tests are "voluntary" in the same sense that the usual attitude test, interest inventory, or neuroticism test is. The respondent is told (either directly or in effect) that there are no right or wrong answers, and he is placed in a situation in which a voluntary and arbitrary performance is acceptable.

DISGUISED-STRUCTURED TESTS OF SOCIAL ATTITUDES

The approaches which we will consider below differ from the ones mentioned above perhaps only in degree or in relative emphasis. Yet the distinction involved is important. The characteristics of these disguised, non-voluntary tests, can be stated in a number of ways:

> The respondent participates in an objective task, in which he seeks right answers. The voluntarism of the usual projective techniques is lacking. To the respondent, the situation is similar to that of an achievement or ability test.
>
> All respondents have a common motivation in taking the test. All, we may assume, are seeking to perform adequately on the same objective task. Attention is focused on a common goal, oblique to the experimenter's purpose.
>
> Rather than stressing the freedom and lack of structuring, there is an attempt to diagnose attitudes from systematic bias in the performance of an objective task. The test may be highly structured—directly scorable—and still offer opportunity for unconsciously operating bias to distort behavior in a systematic and diagnosable manner.

Here is a simple formula for constructing such a test. Find a task which all your respondents will take as objective, and in which all will strive to do well. Make the task sufficiently difficult, or use a content area in which the respondents have had little experience or opportunity for reality testing. Load the test with content relative to the attitude you study. Look among the responses for systematic error, or for any persistent selectivity of performance. If such be found, it seems an adequate basis for the inference of an attitude.

To be sure, this approach has to some extent been used in the projective techniques discussed above. The T.A.T., for example, is introduced as a "test of creative imagination" (**58**), but little attempt is made to carry out effectively this pretense pattern. Respondents accept the task with a wide diversity of motivations, varying levels of seriousness, curiosity, and suspiciousness. And while all may be "projecting," the motivational situation is highly un-uniform, and meaningful comparisons between the responses of the serious, the suspicious, and the playful can hardly be made. Similarly, the "play technique" has been given to college men with the excuse that the experimenter was interested in "ideas for moving picture plays" and wanted the respondent to construct a "dramatic scene" (**58**, p. 553); and the sentence completion test has been given as an ability test (as Ebbinghaus intended it), with instructions to complete the sentences in any grammatical fashion as quickly as possible. But the disguise in these typical projective tests seems to be half-hearted, or implausible, weakened by the explicit freedom of response allowed. At best, the respondents accept the task as meaningless, or as a psychologist's mystery. It is probably common that something of the intent of the experimenter is divined, and the task completed despite this.

Information tests. To make the model of the indirect, disguised, non-voluntary test of social attitudes more explicit, let us examine the use of an information test for attitude diagnosis. The reader has probably commented at one time or another upon the inextricable interrelationship between people's attitudes and what they take to be facts. You would all probably agree that in a detailed test of information, the direction of people's guesses or misconceptions will frequently bear a relationship to their attitudes. In a complementary fashion, a given person's knowledge is apt to reflect in its unevenness his selective awareness and retention, or his biased sources of information. Newcomb has dramatically portrayed the non-random character of right and wrong answers on an information test in his study of "The Effect of Social Climate upon some Determinants of Information" (**62**). While selective awareness and divergent sources of information were primarily involved, Newcomb comments with regard to the difficult information items "the direction of guessing is altogether likely to be weighted toward the subject's attitude. If this reasoning is correct, the . . . test tends to become itself an attitude test." Coffin (**14**) and Smith (**76**) have found beliefs of factual type statements to correlate highly with related attitude tests. Indeed, in looking over the high correlations found in the literature on the relationship between information and attitudes, one is tempted to guess that some re-interpretation of them is in order. Is it not likely that many items were such that persons of liberal, tolerant attitudes found guessing easier?

Loeblowitz-Lennard and Riessman have proposed that an information test be used to measure indirectly attitudes toward Negroes and Jews (**50**). Hammond (**31**) has actually used such a test to measure attitudes toward "labor-management" and "Russia," and his work is worthy of some detailed comment here. We have mentioned that not only guessing behavior but also differential patterns of information may be diagnostic of attitudes. Hammond eliminated the latter by the "error-choice" technique, in which the respondent was forced to choose between two alternative answers each of which was, by intent, equally wrong, but in opposite directions from the correct answer. Such items were:

"Average weekly wage of the war worker in 1945 was (1) $37.00, (2) $57.00." "Financial reports show that out of every dollar (1) 16¢, (2) 3¢, is profit." Scoring these information guesses as attitude items gave total scores on 20 such items that differentiated a labor union group from two business clubs with almost no overlap. In spite of the small number of cases (18 union and 42 business) the critical ratios were 11.3 on the Labor-Management, and 12.5 on the Russia questionnaire. Reliabilities on the two scales were roughly estimated at .78 and .87 respectively.

What we have here is an objective test situation (which could be made more so by using more alternatives and including a correct one) in which people's errors are not random, but systematic. The presence of biased performance clearly necessitates the inference of some underlying process, which we choose to call attitude. The claim for face validity on such a test seems to be stronger than the one that can be made for either the direct or unstructured attitude test. Is not systematically biased performance in dealing with environmental actualities the essential practical meaning of attitude?

Parrish (**64**) has more recently used a regular multiple choice information test as an indirect indicator of attitudes, in conjunction with a direct attitude scale and a second indirect test involving estimating opinion poll results (see below). All three tests effectively reflected 12 minutes worth of anti-Kuomintang propaganda in approximately the same degree when the greater reliability of the direct test is considered. The pseudo-information test had 32 items and a reliability of .66 (using Kuder-Richardson formula 14 modified for multiple levels of response). It correlated .67 (corrected for attenuation) with the direct test (10 items, reliability .91) given at the same time (both anonymously, so there is no reason to suspect the direct test). The correlation with the public opinion estimate test was .59.

Kremen (**43**) attempted to evaluate the effect of student role-playing of a discrimination episode upon attitudes toward the Negro, using a multiple-choice information-type indirect test and a direct test. While neither test reflected the role playing in mean scores, her findings have importance for attitude measurement, inasmuch as role-playing *lowered* the relationship between the direct and indirect test. The results from ten sociology classes, (five pairs in which the instructor and subject were the same) are presented below:

	Correlation between Direct and Indirect Tests				
Role-Playing Classes:	.39	.36	.20	.53	.38
Paired Control Classes:	.54	.70	.35	.58	.60

For all classes combined, the direct test had a reliability of .89, the indirect test .42 (Kuder-Richardson, Formula 14). The explanation of this phenomenon is not obvious from her data insofar as analysed. She does not report reliability values separately for the ten classes, although this step is planned in further analysis of the data.

R. B. Cattell (**12, 13**) has undertaken a large scale exploration into the "objective" measurement of attitudes, using a rationale very similar to the one presented here for non-voluntary attitude tests. In this study, a large number of ingenious methods have been tried, and others suggested.[1] The methods have

---

[1] Omitted from the present review are a number of physiological and related psychological measures, such as P.G.R., writing pressure, pulse pressure, speed of reading, fluency, and speed of decision, which Cattell also investigated.

8-358
7.9.70

been applied to a number of attitude-interests, such as playing more sociable games, excelling in one's career, and being smartly dressed, which differ greatly from the social-problem topics of most of the other studies reviewed. Two of the methods may be classified with information test approaches. His measure of "False Belief (Delusion)" (**12**) was presented to respondents as a multiple-choice information test. A sample item: "During the war church attendance increased greatly and, since V-J day it has: (declined slightly; tended to increase still more; stayed at its high peak; returned to its pre-war level; fallen to its lowest point since 1920)." Presumably the different responses were given different weights as in a direct test and as done by Parrish and by Kremen. Applied to a religious attitude, ten such items gave a Spearman-Brown reliability of .53, and correlated .33 with a paired-comparisons preference test, and .10 with records kept of time and money expenditure.

Cattell's other information test approach was more straight-forward. As mentioned above, we may expect that a person will reveal his attitudes not only in the direction of his "errors" but also in the selective character of the "right answers" he knows.

Cattell prepared 10-item information tests for each of a number of attitudes. The items dealt with the "knowledge required in following the course of action connected with the attitude" (**12**). For twelve topics, these tests, scored in terms of the number of items correct, provided Spearman-Brown reliabilities ranging from .13 to .92, with an average of .59. Correlation with records of time expended on the various activities averaged but .16. With money spent on activities the average value was .15, and with paired-comparison preferences, .16. In this approach, Cattell has resurrected an approach to the "objective" or "indirect" measurement of vocational interests which flourished during and after the first world war. Fryer (**26**) and others provide summaries of such research. No longer called "interest" tests, these selective vocabulary and object recognition tests are today firmly established among vocational selection and guidance tools.

Estimation of group opinion and social norms. As Travers (**79**), Wallen (**80**), and others have demonstrated, there is persistent correlation between a person's own attitude and his estimate of group opinion. While a few persons may chronically underestimate the popularity of their own opinions, the prevailing tendency is to overestimate the size of the group agreeing with one's self. Biserial correlations between endorsement of a dichotomous item and the estimation of the proportion of the group endorsing show few values below .2 and characteristically run from .4 to .6 or higher. This relationship seems to indicate the possibility of assessing a person's attitudes by giving him the entirely objective task of estimating group opinion. If one regards correlation with direct attitude expression as validation, one starts with very high single item validities indeed. If from a single item, one can predict a person's attitudes to the extent of a correlation of .40, how well might one not be able to predict from twenty such items? In several attempts along this line at the Ohio State University, conducted under the direction of the present author, the results have fallen considerably below such optimistic hopes. From combined estimates of group opinion, the best correlation obtained with a direct attitude score based on the same items runs around .60, not better than the best single item-estimate correlation. More detailed exploration of the technique is being made, however. Parrish (**64**) used a public opinion estimate test, with items dealing with U. S. opinion on China, the opinions of Americans with experience in China, and the opinions of the Chinese. Only the first type are directly comparable to the Travers and Wallen situation, and these did not come out the best in the item analyses. The test, containing only 10 items, had a Kuder-Richardson reliability of .53, correlating .67 with the direct test when corrected for attenuation. In this instance there was no direct duplication of content in the direct attitude test items and in the opinion estimates.

The technique is not entirely new. In 1929, Sweet, using a test designed by Goodwin Watson, found boys' estimates of group opinions valuable in diagnosing adjustment problems (**77**). More recently, Katz has reported results confirming the general relationship between own attitudes and estimates of group opinion, although he does not give correlation coefficients (**41**). As part of Section A in their "Sentiments" examination, Murray and Morgan (**59**) ask the respondent to guess what the majority of people believes or prefers, e.g. "what are the three most popular things to do?" Newcomb (**61**) made one of the first uses of the percentage estimate of group opinion, but has not reported his results in terms of the responses of individuals. It might be pointed out that the correlation involved in these studies is not necessarily to be interpreted as projection. In a group of any size, the respondent has an uneven acquaintanceship, and probably associates more with those of tastes like his own. Basing his estimates of group opinion upon his own experience in the group, his error may be in part a "sampling bias" as well as biased perception.

F. C. Bartlett (**4, 5**) and D. M. Carmichael (**9**) before the war were experimenting with the use of predictions of social events as diagnostic of attitudes. The respondent is given a situation and some background information, and then asked to make a prediction. Predictions regarding the likelihood of cooperation between social groups have been used, for example. The approach can probably be presented as a test of ability, and would seem to offer the possibility of wide variation in content. McGregor (**53**) has also pointed out the relationship between predictions and personal attitudes. Using predictions of another kind, Davis (**17**) has devised a simple test for use with children which could easily be given in an objective frame of reference, and which already has the desirable quality of asking the respondent to talk about somebody else rather than himself. (Indeed, his technique is practically identical with the first "projective" personality tests designed by Cattell (**10**) on the model of multiple-choice intelligence test items.) In Davis' test, situations are presented, and the respondent predicts the protagonist's action by selecting from several prepared alternatives.

Noland (**63**) used a projective approach toward estimating social norms with one set of items in his extensive study of absenteeism, asking "how important a cause of absenteeism do you think each of the following." These single items predicted absenteeism with a correlation of .50 or better. The value of the finding is weakened, however, by having as a criterion of absenteeism voluntary reports on the same (unsigned) questionnaire.

Tests of ability to do critical thinking. During the last three or four years of his life, John J. B. Morgan had been working on another technique that admirably meets our criteria (**55, 56, 57**). He used a syllogisms test, with a large number of possible conclusions provided, from which the respondent was to pick the logically correct one. Following Sells' study of atmosphere effect, the syllogisms were invalid. In spite of the invalidity of the syllogism, college students will concentrate their choices on the one or two wrong alternatives favored by the atmosphere effect. To use this phenomenon in diagnosing attitudes, Morgan gave the same syllogisms in both impersonal or abstract form (e.g., "No A's are B's. Some C's are B's. From these statements it is logical to conclude: No C's are A's, etc.") and with content ("A trustworthy man does not engage in deceitful acts. The bombing of Pearl Harbor by the Japanese was a deceitful act. From these statements it is logical to conclude: No Japanese are trustworthy, . . . etc."). By studying the shift in the popular responses from the nonsense to the meaningful form, Morgan attempted to diagnose group attitudes. While the technique needs much further study, it seems well worth such elaboration. In particular, an attempt should be made to control the effect of greater tangibility, per se, and also the intrusion of evidence from outside the syllogism which might be confounded with purely attitudinal effects. Note that in this particular study, the shifts of the group showed more tolerance for the Japanese than otherwise. Lansdell has utilized the technique for studying covert attitudes toward marriage, attempting to apply it for the diagnosis of the attitudes of a single individual (**45**).

Ruch, in his 1937 edition of *Psychology and Life* (**69**, pp. 633–634), reports briefly upon a classroom experiment in which preferences for ethnic groups were reflected in the ability to judge the validity of syllogisms about the same groups. In this instance, no choice of conclusions was offered, but rather the students judged the validity of the complete syllogism as stated. Lefford (**46**) has shown in more detail the disrupting influence of emotional subject matter upon such judgments.

The syllogism approach, and a number of others, have been anticipated by Goodwin Watson's *Measurement of Fair Mindedness*, published in 1925 (**81**). By fair-mindedness Watson seems to have meant something roughly equivalent to tolerance, critical thinking, or open-mindedness. His six sub-tests were designed to provide scores on this trait, which they did with adequate reliability. But Watson recognized that in the "errors" or lapses from fair-mindedness were clues as to the biases or attitudes of the respondents, and a secondary "analytic" scoring was made in these terms. Three of the sub-tests we would characterize as being direct[2] and of the others, we will classify two under the heading of "critical thinking," treating the third—*Moral Judgments*—below. The *Inference Test* provided statements of fact, followed by several conclusions that might be drawn, only one of which is logically justified, the others offering opportunties for the intrusion of personal biases in various directions. This test has been expanded in the Watson-Glazer Tests of Critical Thinking (**28**) which could be scored for various prejudices although they have not been so used as far as the present writer is aware. Gilbert (**27**) has made use of this same technique in a study in which high-school children were given hypothetical racial problems for which they could choose strictly logical conclusions or conclusions showing bias.

---

[2] These are the *Word Cross-out Test* in which the respondent crosses out "distasteful" words, as in Pressey's test, the *Degree of Truth Test* allowing for extreme or moderate statements of belief and disbelief and, the *Generalization Test* in which a number of statements about various groups are presented, and the respondent is to indicate whether the statement is true of "All, Most, Many, Few, or No" members of the group. (This is probably one of the first tests of "stereotyping." It might be classified as indirect if administered as an information test.)

In Watson's *Arguments Test* both pro and con arguments on various topics are presented, and the respondent evaluates these for their strength as arguments. The test taps the tendency for a person to view as strong those arguments which support his own point of view. Robbins (**66**) has recently confirmed the principles underlying this test, although his data are not in such form as to indicate whether or not accurate inferences as to attitudes could be made from such judgments. He found that students performed better in judging the adequacy of reasons for points of view with which they agreed than those with which they disagreed.

Tests employing bias in perception and memory. The demonstrations of systematic attitudinal bias in highly "objective" perceptual and learning tasks have provided some of the most important developments of the past 15 years for attitude theory (**2**, **20**, **48**, **73**, **74**). The use of errors in seeing and remembering in the diagnosis of individual attitudes epitomizes the disguised, structured, and non-voluntary assessment of attitudes.

Pioneers in the use of this approach are E. L. and R. E. Horowitz (**37**). In their study of the development of social attitudes among the children of a southern community they invented a number of techniques which have merited a great deal more re-use than they have seen. While designed to portray group differences, many of them should be appropriate to individual testing. Their *Aussage Test* involved exposing a complicated picture for two or three seconds, following which they tested for perception and memory through a series of standardized deliberately leading questions. For instance, the question "who is cleaning the grounds" to Picture 10 brings answers referring to a non-existent Negro in some 70% of the cases. To Picture 4, the misleading question "What is the colored man in the corner doing" brings a steadily increasing proportion (in reports from higher age groups) of menial activities for the non-existent Negro man.

Their *Perception-Span Test* involved a series of posters each having pictures of 10 items mounted upon it. These were exposed for 10 seconds, and the children asked to "tell all the pictures you can remember." While there were difficulties with perseveration and blocking, some age group differences were noted, with the younger children failing to note Negroes, and the older ones showing a selective awareness for them. A *Recall Test*, asking for the reproduction of ten words used earlier in a word association test, seemed also to indicate that in the younger years the word "Negro" was less well remembered than would be expected. The *Pictorial Recognition Test* involving the recognition of faces previously exposed, finds the faces of Whites more frequently recognized than the faces of Negroes.

Seeleman's study (**71**) confirms this finding with respect to individual differences within the group. Presenting sets of Negro and white photographs she tested for later recognition by asking that the previously exposed pictures be picked from a larger group. The discrepancy between memory for white and Negro faces was checked against attitudes on a direct test. Using people in the extreme quartiles on the direct test only, the test and memory-bias scores correlated .64 and .71 with different populations. In a paired-associates memory test involving complimentary phrases to be paired with white and Negro pictures, error discrepancy scores showed no correlation with the direct test, but the tendency to assign favorable phrases to the Negro (regardless of correctness) correlated .66 with the direct test in the one population for which the correlation is reported. These values are high enough to suggest that with refinement an indirect test might be achieved.

On a preliminary experiment, Murray Jarvik (22) attempted to utilize memory distortion as an indicator of attitudes. Sixth, seventh and eighth grade students in a rural school were read five minute stories and then were asked to write down all they could remember of the story. The stories had simple dramatic plots, but were full of confusing detail and lacunae with regard to specific names, characteristics, and ethnic identities. Opportunity was given for memory distortion in the direction of common stereotypes. The results were essentially negative. However, motivation and literacy were low, and the task was a part of a rather long testing program. Note also that memory distortion usually occurs over longer spans of time, and that had the test employed selective retention rather than distortion it might have been more successful.

Cattell (12, 13) utilized three techniques involving memory and perceptual distortion. His test of "Immediate Memory" was based upon selective recall for attitude relevant statements from sets of twelve presented at one-second intervals. In the total, there were over 500 statements. Spearman-Brown reliability figures for eleven attitudes ranged from .13 to .86, averaging .50. Correlations with other measures were essentially zero (with time expenditure, average .01; with money expenditure, average .03; with paired-comparisons preferences, average .13; with the information test (see above), .01.). In the scoring, both "facilitating" and "frustrating" statements relative to the attitude were pooled. Cattell recommends that in future research these be separated.

Cattell's "Distraction" method involved the exposure for ten seconds of statements related to the attitudes. Around the statement were scattered twelve or thirteen nonsense syllables. Subjects were held responsible for recalling the statement and the nonsense syllables. Scoring was in terms of decrement in nonsense syllable recall for statements related to a given attitude. The original hypothesis was that the stronger the attitude, the poorer the nonsense syllable learning, through distraction. Actually, the opposite effect was found. With a Spearman-Brown reliability of .64 for one attitude (being smartly dressed), the correlation with the paired-comparison preference measure was −.29, with time expenditure, −.10, with money expenditure, −.08, and with the information test (discussed above) −.35. Since only ten items were involved, these values are regarded as sufficiently high to encourage further work with the method. The "Misperception" method involved the tachistoscopic exposure of attitude statements for one second, with the respondent responsible for repeating the statement and noting the misspellings. The overlooking of misspelling was regarded as a sign of strong attitude. Ten items dealing with the desire to play more sociable games provided a Spearman-Brown reliability of .43, and correlations with the paired-comparisons preference measure of .00, the expenditure of money of −.01, and the information test of −.13.

These methods of selective memory and selective distractibility have been anticipated in the indirect measurement of occupational interests by Burtt in 1923 (**7**, **8**, pp. 334–343). In a paired-associates memory test, one half of the pairs were relevant to agricultural engineering. The ratio of memory for these pairs as opposed to memory for other pairs correlated .30 with instructors ratings as to interest and industry. In a cancellation test, in which irrelevant words were to be crossed out, different paragraph contents were employed. Differentials in efficiency of cancellation correlated .30 with the above-mentioned instructors ratings. Contrary to Cattell's findings, the greater distractibility was associated with stronger interest.

Tests involving ability to judge character. Murphy and Likert, in *Public Opinion and the Individual* (**60**), utilized a wealth of techniques anticipating the projective testing movement. Among these was the "Ratings from Photographs." Following the general framework of Rice's classic study on stereotypes, they provided labeled pictures of a union president, a railroad magnate, a pacifist, a Negro champion of Negro rights, etc. Respondents were asked to judge from these labeled photographs the character of the pictured person, in terms of courage, selfishness, intelligence, conceit, sympathy, practicality, and sentimentality. Contrary to expectation, they found no relationship between attitudes as measured in a variety of paper and pencil tests and these picture ratings. However, since the making of character judgments from photographs comes as near to being a paradigm of psychiatric projection as can be found in present test approaches, it might be suggested that this technique be further explored before being abandoned. In line with the tenor of the present discussion would be the suggestion that the objectivity or "pseudo-objectivity" of the task be increased, with the use of some such label as "social intelligence test" or the like.

In precisely this model is the "Faces Game" originally developed by Marion Radke (**67**, p. 50–51) and being used as modified by Isidor Chein and Iljana Schreiber in research of the Commission on Community Interrelations of the American Jewish Congress (**16**). In this test children are asked "Let's see how good you are at telling what people are like just by looking at their faces." Thirty four sets of four pictures (two white and two Negro for all but four sets) are presented, each with a question such as "one of the girls in this row is very lazy and never bothers to do anything" or "one of the boys in this row is the best sport in his class." Two scores are provided: one *the "white salience" score* (the total mentions of white children for items either good or bad), and *the prejudice score* (the ratio of unfavorable choices of faces of the other group to total choices of other groups). The test produces significant individual and group differences, and showed retest reliabilities over a six month period of .50 and .36 for white and Negro children, respectively, on the prejudice score. For the white salience score, the values were .32 and .16. The researchers do not regard the test as satisfactory for general use as yet.

Hsü (**38**) had three women graduate students sort male photographs for handsomeness, and ten days later for judged membership in the Communist party. The correlations were negative (−.50 and −.21) for the two women who were anti-communist and positive (.51) for the one woman who was relatively pro-communist. After reading a *Time* report on the blockade of Berlin, the correlations for sorts on a second set of the photographs were negative for all three women. While this study does not provide in itself an adequate basis for measuring attitudes, it does hold out promise for more extensive work with character judgments from photographs.

Tests involving miscellaneous abilities. Another of the excellent test approaches in the Horowitzs' study (**37**) was labeled the *Categories Test.* This test was modeled directly upon a typical intelligence test item. Five pictures were presented, and the question asked "Which one does not belong?" As they adapted it, the item could be answered in more than one way, and the choice indicate something of the important social categories for the child. For example, a page might contain pictures of three white boys, one white girl, and one negro boy. In this instance, the child could use either sex or race as the dominant category. Other items provided opportunities to categorize by age and socio-economic status as well. Here again, the emphasis was on group norms rather than individual differences.

Hartley (**32**) has more recently used the categorization of photographs as a measure of ethnic salience. In this study, however, the task was presented in a voluntary framework, in contrast with the objective assignment presented in the *Categories Test.* In Hartley's "faces test" the respondent is told to: "Sort them out into piles. You can make as many piles as you want to, or as few as you choose. You can classify them on any basis you want to!"

In the Murray and Morgan "Study of Sentiments" (**59**) several of the "indirect methods, methods which conceal the examiner's intent" provide an ability test façade. The "Sentiments Examination" Section A Part 1 asks for the "most descriptive adjectives" for 48 stimulus words, "the S being led to believe that the examiner is interested in testing the range of his vocabulary." In Part 3 "The S is led to believe his verbal ability is being tested with a simile completion test," e.g. "As pathetic as a . . . ." In the "Arguments Completion Test" the respondent is asked to continue and finish an argument, the beginning of which has been described. "Being led to believe his powers of argumentation are being tested, the subject quickly becomes involved in the controversy he or she is inventing, and ends by exposing more of his own sentiments than he might otherwise have done" (**59**, pp. 58–60).

Tests involving miscellaneous judgments. The following tests are disguised and structured, but in part at least must be classified as *voluntary*, in so far as judgments are required in situations in which there is no "objective" right answer. On the other hand, they retain the advantage of having the respondent (1) work on a task presumably less threatening than the experimenter's primary problem, and (2) report upon external values or realities, rather than upon himself directly. These studies like many already discussed, depend upon the demonstration of a judgment differential of which the respondent is presumably unaware.

In Watson's (**81**) "Moral Judgments" sub-test, judgments of approval or disapproval are made about a variety of situations, sets of these situations being identical except for the specific persons or groups involved. For example unwarranted search is made of a suspected "radical" headquarters, on the one hand, while in a parallel item, the same type of search is carried out with a business corporation suspected of dishonesty. Scoring is done on the basis of discrepancy of judgment between the parallel situations.

Similar in title and in plan to Watson's Moral Judgments Test is the ingenious approach to Negro-White attitudes devised by Seeman (**72**). While he used equated comparison groups, the technique could be modified for diagnosing bias in individuals. He selected six items from a standard test of moral evaluations on marriage and sexual matters. To one group, the episode items were illustrated by pictures of white couples; to the other, with Negro illustrations. Both groups were made up of white college students. Contrary to expectation, the judgments were more lenient—less disapproval—with the Negro illustrations. Furthermore, when the two groups were sub-classified according to scores on the Likert scale of attitudes toward the Negro, the major part of the differential was contributed by the more *tolerant* persons rather than the more anti-Negro parts of the two groups. The intolerant extremes in this sample were more consistent, less "biased." These results are important and meaningful, but further indicate the danger of over-simplified interpretations in indirect approaches.

The readiness with which judgments of literary merit can be manipulated by the substitution of fictitious authors has been demonstrated by Farnsworth (**70**) and Sherif (**73**). Prestige seems to be one of those forces which can bias the performance of an objective task. Can we use this bias to infer attitudes? Lewis (**47**) seems to have had in mind some such approach. The present writer (**22**) deliberately set out to utilize the prestige effect to measure attitude toward five minority groups. Rather than literary passages, proverbs or mottoes were used. Rather than individual authors, the adage was attributed to the group as a whole, e.g., "American pioneer saying" or "old Jewish motto." Eight and ninth grade students were asked to evaluate the quality of each motto separately—there being ten mottoes attributed to each group. The test yielded a general prejudice score that correlated but .30 with a direct test. Other scores were worthless. However, in view of the strength of prestige suggestion in other experiments, this approach seems worth trying further. It is quite possible that in the form given the task was trivial and the disguise thin.

Horowitz (**35**) illustrated the development of patriotism and conformity among children by tests in which the task was to pick "the best looking flag," or the "best place to live in" from a series of photographs of houses, etc.

As Wolff, Smith and Murray (**83**) have shown, reactions to group-disparagment jokes are correlated with group membership. Gordon (**29**) has recently tried to utilize this phenomenon in the assessment of social attitudes. Twenty four jokes, both antagonistic and sympathetic, dealing with Negroes and Jews were rated as to their funnyness on a 5-point scale. Four groups of college men were used: a Protestant fraternity, a Catholic fraternity, a Negro fraternity, a Jewish (non-Zionist) fraternity, and a Zionist club. The groups differed in their responses to these jokes but not as anticipated in all instances. The jokes sympathetic to Jews were rated highest by the two Jewish groups—but the anti-Jewish jokes failed to differentiate the groups. The anti-Negro jokes were rated lowest by the Negro group, but the pro-Negro jokes failed to differentiate. The various groups ranked the individual jokes quite similarly as to popularity, with rho's between the orders for the different groups ranging from .71 to .94. With regard to individual differences *within* the various groups, the ratings of the jokes showed no relationship to attitudes toward Negroes or Jews as revealed in a direct attitudes test or a symbol endorsement test. The tests were administered as a part of a larger testing program.

Cattell (**13**), in the program already referred to used a "Projection-Phantasy" test, in which the respondent picked from ten sets of ten incomplete sentences, five each which he completed. Rather than scoring the words used in completing the sentences, Cattell scored the *choice* of incomplete sentences. Each set contained one representative of each of ten different attitudes. Applied to the desire to excel in one's profession, and to the desire for more sleep and rest, this test had Spearman-Brown reliabilities of .13 and .60. Correlations with the paired-comparison preferences were .15 and .31, with time and money expenditure, .01 and .26, with information, .20 and .20. In spite of the relatively low level of these values, this method was regarded as one of the more promising.

Kalpakian's "The Construction of a Disguised Test by the Use of Photographs for the Study of Attitudes Toward Negroes" (40) represents an attempt at disguise which fails in part to meet the criteria suggested here. Like other tests reported in this section, a judgment differential of which the respondent was unaware was sought. Twenty-six matched pairs of pictures were selected—one set depicting Negroes, the other whites. The test was structured, in that a choice of responses, "Like—Indifferent—Dislike" was required for each picture. The instructions read in part "we're interested in seeing what kinds of pictures people like and don't like . . . we're not interested in the technical photographic aspects of the pictures, rather we want your reactions to what is in the pictures. That is, how do you feel about what is in each picture." Note the voluntary nature of this assignment—the respondent being told in effect that there are no right or wrong answers. Note also the absence of any clear cut assigned basis of judgment which might lend indirection (such as asking for judgments of beauty of the photograph or skill of photographer). Kalpakian did aid the disguise by placing these 52 photographs in random order among 68 irrelevant pictures. This is similar to the efforts in disguise made by those who distribute the items of one attitude scale among a variety of other types (e.g., 24). In one of Kalpakian's groups, 100 subjects were asked what they thought the purpose of the test was. Only 15 did not guess that the purpose involved attitudes toward the Negro; 51 thought this and several other attitudes were being tested; 34 judged correctly that the primary purpose was attitudes toward Negroes. On the other hand only 7 were aware of the matching of the pictures. Scored on proportion of white to Negro "likes" the test showed corrected split-half reliability coefficients between .71 and .82 in various white high school and college groups. Mean values showed a definite "white salience" for all groups. Scores correlated with both self ratings and ratings by friends, demonstrated by critical-ratio tests. The correlation with the Thurstone Hinkley was .40, which became .80 when corrected for restricted range. In terms of the categories of this review, Kalpakian's test is similar to the "show me" and other picture tests developed by Horowitz (34). It is structured in that a limited number of response categories is provided. It is direct, substituting pictures for the ethnic labels used in statement-endorsement type of direct tests.

GENERAL CONSIDERATIONS

Validity of indirect tests. Most efforts to develop indirect attitude tests are predicated upon the assumption that indirect tests will under certain conditions have higher validity than direct tests. It is worthy of note that *none* of the studies reviewed here offers any evidence that this is the case. The testing of this hypothesis is certainly the next step if these preliminary efforts are thought to be sufficiently promising to justify more research.

The most usual effort at validation has involved correlating the indirect test with direct test results, under conditions in which the latter were gathered anonymously or in which there was no reason to suspect respondent dissembling (e.g., **19**, **22**, **29**, **38**, **40**, **43**, **60**, **64**, **65**). Hammond (**31**) and Gordon (**29**) used environmentally selected criterion or comparison groups. Parrish's tests (**64**) meaningfully reflected an experience differential between two groups. These are, however, isolated instances, and even at the exploratory level, we might expect more use of criterion groups.

As has been pointed out previously, the case for *face validity* for indirect tests may often be better than for direct tests. In a number of the disguised, structured tests, the distribution of scores and measures of internal consistency demonstrates unequivocally that non-random, systematic errors, differences in perceptions, etc. exist, of which the respondents are presumably unaware. These systematic unconscious "biases" are well worth study in their own right, and seem to lie close to the functional meaning of attitude.

Attitude measurement as related to attitude definition and theory. Research on social attitudes has been justly criticized for a lack of common definition of the concept, and for a failure to integrate definition and measurement procedures. This diversity of definition has been in odd contrast with the obvious similarity of research procedures. This paradox arises from definitional attempts which confound *explanations* of the phenomena with the process of *pointing* to the phenomena. It is the contention of the present writer that agreement on the implicit operational (or pointing) definition of attitudes is already present. As a tentative formulation the following is offered: *A social attitude is (or is evidenced by) consistency in response to social objects.* If we look at those definitions utilizing concepts of set, or readiness to respond—for example, Allport's (**1**): "An attitude is a mental and neural state of readiness, organized through experience and exerting a directive or dynamic influence upon the individual's response to all objects and situations to which it is related"—and ask for the evidence of a "mental and neural state of readiness," the symptoms of a "directive or dynamic influence," criteria as to the "objects and situations to which it is related," these evidences will be in final analysis consistency or predictability among responses. *An individual's social attitude is a syndrome of response consistency with regard to social objects.* And even those whose behavioristic orientation leads to a rejection of such mentalistic definitions as Allport's—and who would say with Bain (**3**) and Horowitz (**36**, p. 142), "essentially . . . . the attitude must be considered a response rather than a set to respond"—in research practice do not equate *isolated* responses with attitudes, but on the contrary look for the appearance of *response consistencies*. This is dramatically evidenced by Horowitz's (**34**) use of the appearance of consistant differentiated response to photographs of Negro and white children to mark the occurrence of race prejudice in children.

This standpoint provides a basis for the operational delimitation of specific and general attitudes. It likewise demands that if attitude measurement be integrated with definition, internal consistency among the sample responses collected by the test must be demonstrated. To the present writer, the demonstration of a single factor through factor analysis seems ideal, although time consuming. Spearman-Brown and Kuder-Richardson (**44**) reliability measures, Guttman's (**30**) "reproducibility coefficient," Loevinger's (**52**) "homogeniety," and Thurstone's "coefficient of irrelevance" (**78**) are likewise evidences of internal consistency. Item analysis against total scores (**49**, **75**), remains one of the best practical means of improving internal consistency.

Contemporary American theories commonly hypothesize attitudes to be learned, or identify them with other cognitive phenomena such as perception and concept formation. From the operational definition suggested it would be hard to distinguish attitude from habit, which likewise is evidenced by behavior consistency. Indeed, if learning be conceived broadly enough to include the development of any response consistency, attitude formation is identical with learning. Traditionally, learning has been used in those situations in which the response consistencies represented optimal reflections of a simple, uniform environment—where increasing response consistency represented the elimination of "error." But a similar process goes on in "unsolvable" or highly complicated environments, and stereotyped modes of response are developed which are non-optimal reflections of the environment, which do not represent error-reduction. The terms perseveration, set, fixation, position habit, bias, irrational behavior, idiosyncracy, pre-judgment, non-random error, over simplification, bad habit, rigidity, all come to mind to describe this process. Presumably, this latter type of response consistency is a result of the same cognitive processes which under other conditions result in "right answer" learning. It is with these common processes that the concept of attitude is identified, although typically it is most closely associated with the acquisition of "irrational" or non-adaptive response consistencies.

In these terms it may be meaningful to compare direct and indirect efforts at attitude assessment. In both direct and indirect tests enough internal consistency has been demonstrated on some topics to justify speaking of attitudes. Likewise, in numerous instances reported here, considerable correlation has been demonstrated between responses to direct and to indirect tests, justifying the hypothesis that a common attitude lies behind them. Missing is the demonstration that this pattern of consistency extends beyond paper and pencil to the "real-life" situations which are usually in mind when the concept of "attitude" is used. While meaningful research problems exist in regard to patterns of consistency within a universe of paper and pencil responses, the bulk of the researches reviewed are predicated upon the assumption that broader consistencies exist—that overt behavioral manifestations of the attitude can be predicted.

With regard to the level of response consistency, it should be noted that direct tests uniformly have much higher reliability coefficients than do indirect ones, especially when the number of items and time of administration are considered. Of course this consistency is in part conscious, voluntary, and possibly superficial—in contrast to the involuntary "bias" in performance achieved by many indirect tests. In this sense, the indirect tests of the non-voluntary sort do in themselves demonstrate and utilize the dramatic phenomena of attitudinal interference with tasks of learning, perceiving, remembering, and evaluating, such as are summarized by Sherif and Cantril (**74**, pp. 29–91). And it is such consistencies, rather than consistency in voluntary self-description, which give vitality to current usage of the concept of attitude. In this, indirect attitude tests lie closer to contemporary attitude theory than do direct tests.

SUMMARY

The efforts at indirectly measuring social attitudes have been summarized. A distinction has been offered between disguised, non-structured, types of test on the one hand, and disguised, structured, tests on the other. Because of a greater freedom from dependence upon the respondent's voluntary self description, the use of disguised, structured tests has been emphasized.

As can be seen from the survey, there is an abundance of partially tried techniques. Equally good ones can be developed from the general formula: A plausible task, (a) which your respondents will all strive to do well, (b) which is sufficiently difficult or ambiguous to allow individual differences in response, and (c) which can be loaded with content relative to the attitude you seek to measure. Test the responses of individuals for persistent selectivity in performance, for correlated or non-random errors. For example, several persons have suggested to the present writer what might be called the "photocrime" test. Here criminal scenes would be pictured in photograph and text. Sufficient ambiguity would be provided to make it a challenging problem to the amateur detective, and, incidentally, to allow the possibility of bias in the selection of the criminal from the suspects of various ethnic or social identification. A test of "common legal knowledge" might well provide a framework for carrying further some of the notions in Watson's "Moral Judgments" sub-test. The respondent might be asked to decide whether or not civil liberties had been violated in a wide range of specific situations—might be asked to decide between claimant and defendant in a number of civil suits. Here the content could easily be manipulated to get at almost any social attitude one might be interested in. It should not be too difficult to keep the respondent convinced that he dealt solely with matters of the law. We are prone to see as propaganda that with which we disagree. A "propaganda analysis test" should reveal indirectly some people's attitudes. For the build-up, give a review of some principles on how to spot propaganda, how to detect an ulterior motive, a selfish interest group, deceit, exaggeration, etc. Then have the respondent test his skill on a number of argumentative paragraphs, which he evaluates as to the extent and kind of propaganda present. The root of this notion may be found in Watson's "Arguments" sub-test.

Such devices may be multiplied; the reader can no doubt provide better ones. In some of these, the fiction of an objective test situation will be hollow indeed. In others, for example in the estimation of group opinion, we are tapping biased performance in an everyday act of social perception. We are not taking at face value the respondent's own volunteered description of himself. Rather we are studying systematic errors in the respondent's own perceptions, errors of which he himself is not aware. Not only in attitude testing but in personality study in general, distortions of performance in dealing with the environment provide objective evidence of an individual's unique picture of his world. Just as the psychophysicists were long ago forced to introduce *Vexirwersuche*, or catch trials, into introspection—to eliminate impossible negative reaction times and infinitesimal two-point thresholds—so personality testing, whether check-list or free-response, must rise above dependence on voluntary self-description. The development of structured indirect attitude tests is thought to be a step in this direction.

BIBLIOGRAPHY

1. Allport, G. W. Attitudes. Ch. 17 in C. Murchison (Ed.), *A handbook of social psychology.* Worcester, Mass.: Clark University Press, 1935.
2. Allport, G. W., & Postman, L. *The psychology of rumor.* New York: Holt, 1947.
3. Bain, R. An attitude on attitude research. *Amer. J. Sociol.*, 1928, **33**, 940–957.
4. Bartlett, F. C. *The study of society.* London: Kegan Paul, 1939.
5. Bartlett, F. C. The cooperation of social groups: a preliminary report and suggestions. *Occup. Psychol.*, 1938, **12**, 30–42.
6. Brown, J. F. Modification of the Rosenzweig picture frustration test to study hostile interracial attitudes. *J. Psychol.*, 1947, **24**, 247–272.
7. Burtt, H. E. Measuring interests objectively. *Sch. & Soc.*, 1923, **17**, 444–448.
8. Burtt, H. E. *Principles of employment psychology* (Rev. Ed.). New York: Harper, 1942.
9. Carmichael, D. M. The cooperation of social groups. *Brit. J. Psychol.*, 1938, **29**, 206–231; 329–344.
10. Cattell, R. B. *A guide to mental testing.* London: Univ. London Press, 1936.
11. Cattell, R. B. Projection and the design of projective tests of personality. *Charact. and Pers.*, 1944, **12**, 177–194.
12. Cattell, R. B., Heist, A. B., Heist, P. A., & Stewart, R. G. The objective measurement of dynamic traits. (To be published in *Educ. psychol. msmt.*, 1950.)
13. Cattell, R. B., Maxwell, E. F., Light, B. H., & Unger, M. P. The objective measurement of attitudes. (To be published in the *Brit. J. Psychol.*, **50**, 1950.)
14. Coffin, T. E. Some conditions of suggestion and suggestibility: a study of certain attitudinal and situational factors influencing the process of suggestion. *Psychol. Monogr.* 1941, **4**, No. 241.
15. College Study in Intergroup Relations. (Lloyd Allen Cook, Director.) Study forms and technics in intergroup relations. Supplementary Sheet No. 5, Jan., 1948. (Mimeographed.) Detroit: Wayne University.
16. Commission on Community Interrelations of the American Jewish Congress. (Stuart Cook, Research Director.) The face game. Unpublished research on a modified version of a test originally developed by Marian Radke. 1947–1948.
17. Davis, T. E. Some racial attitudes of Negro college and grade school students. *J. Negro Educ.*, 1937, **6**, 157–165.
18. Deri, S., Dinnerstein, D., Harding, J., & Pepitone, A. D. Techniques for the diganosis and measurement of intergroup attitudes and behavior. *Psychol. Bull.*, 1948, **45**, 248–271.
19. Dubin, S. S. Verbal attitude scores predicted from responses in a projective technique. *Sociometry*, 1940, **3**, 24–28.
20. Edwards, A. L. Political frames of references as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1948, **36**, 34–61.
21. Evans, M. C., & Chein, I. The movie story game: a projective test of interracial attitudes for use with Negro and white children. Paper read at the 56th annual meeting of the American Psychological Association, Boston, Sept. 8, 1948.
22. Frenkel-Brunswik, E., Jones, H. E. (Directors), & Rokeach, M., Jarvik, M., & Campbell, D. T. (Staff). Unpublished research on the personality correlates of anti-mi-

nority attitudes among grade school children. A project of the University of California Institute of Child Welfare, financed by a grant from the American Jewish Committee, 1946–1947.

23. Frenkel-Brunswik, E., & Reichert, S. Personality and prejudice in women. (Unpublished manuscript.)

24. Frenkel-Brunswik, E., Levinson, D., & Sanford, R. N. The antidemocratic personality. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology.* New York: Holt, 1947.

25. Fromme, A. On use of qualitative methods of attitude research. *J. soc. Psychol.*, 1941, **13**, 429–460.

26. Fryer, D. *The measurement of interests.* New York: Holt, 1931.

27. Gilbert, H. H. Secondary science and pupil prejudice. *J. educ. res.* 1941, **35**, 294–299

28. Glazer, E. M. An experiment in the development of critical thinking. *Teach. Coll. Contr. Educ.*, No. 843. New York: Bureau of Publications, Teachers Coll., Columbia Univ., 1941.

29. Gordon, S. Exploration of social attitudes through humor. Master's Thesis, Univ. of Illinois, 1947.

30. Guttman, L. A basis for scaling qualitative data. *Amer. sociol. Rev.*, 1944, **9**, 139–150.

31. Hammond, K. R. Measuring attitudes by error-choice; an indirect method. *J. abnorm. soc. Psychol.*, 1948, **43**, 38–48.

32. Hartley, E. L. *Problems in prejudice.* New York: Kings Crown Press, 1946.

33. Hartley, E. L., & Schwartz, S. A pictorial-doll play approach for the study of children's intergroup attitudes. Mimeographed preliminary draft. Research Institute in American Jewish Education, American Jewish Committee, Summer, 1948.

34. Horowitz, E. L. The development of attitude toward the Negro. *Arch. Psychol., N. Y.*, 1936, **28**, No. 194.

35. Horowitz, E. L. Some aspects of the development of patriotism in children. *Sociometry*, 1940, **3**, 329–341.

36. Horowitz, E. L. 'Race' attitudes. In O. Klineberg (Ed.), *Characteristics of the American Negro.* New York: Harper, 1944.

37. Horowitz, E. L., & Horowitz, Ruth E. Development of social attitudes in children. *Sociometry*, 1938, **1**, 301-338.

38. Hsü, E. H. An experimental study of rationalization. *J. abnorm. soc. Psychol.*, 1949, **44**, 277–278.

39. Johnson, G. G. An experimental analysis of the origin and development of racial attitudes with special emphasis on the role of bilingualism. Ph.D. Thesis, Univ. of Colorado, 1949.

40. Kalpakian, E. Y. The construction of a disguised test by use of photographs for the study of attitudes toward Negroes. Master's Thesis, Clark University, 1947. (*Clark University Bulletin.* Abstracts of Dissertations and Theses, Worcester, Mass., Oct., 1947.)

41. Katz, Martin R. A hypothesis on anti-Negro prejudice. *Amer. J. Sociol.*, 1947, **53**, 100–104.

42. Krech, D., & Crutchfield, R. S. *Theory and problems of social psychology*, New York: McGraw-Hill, 1948.

43. Kremen, E. O. An attempt to ameliorate hostility toward the Negro through role playing. Master's Thesis, The Ohio State University, 1949.

44. Kuder, G. F., & Richardson, M. W. The theory of the estimation of test reliability. *Psychometrika.*, 1937, **2**, 151–160.

45. Landsdell, H. A study of distorted syllogistic reasoning as a means of discovering covert attitudes toward marriage. *Bull. Canad. Psychol.*

Assn., 1946, **6**, 98. (Abstract.)

46. Lefford, A. The influence of emotional subject matter on logical reasoning. *J. gen. Psychol.*, 1946, **34**, 127–151.
47. Lewis, H. B. An approach to attitude measurement. *Psychol. League J.*, 1938, **2**, 64–67.
48. Levine, J. M., & Murphy, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, **38**, 507–517.
49. Likert, R. A technique for the measurement of attitudes. *Arch. Psychol., N. Y.*, No. 140, 1932.
50. Loeblowitz-Lennard, H., & Riessman, F., Jr. A proposed projective attitude test. *Psychiatry*, 1946, **9**, 67–68.
51. Loeblowitz-Lennard, H., & Riessman, F., Jr. A preliminary report on social perception test—a new approach to attitude research. *Social Forces*, 1946, **24**, 423–427.
52. Loevinger, J. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol. Bull.*, 1948, **45**, 507–529.
53. McGregor, D. The major determinants of the prediction of social events. *J. abnorm. soc. Psychol.*, 1938, **33**, 179–204.
54. McNemar, Q. Opinion-attitude methodology. *Psychol. Bull.*, 1946, **43**, 289–374.
55. Morgan, J. J. B. Distorted reasoning as an index of public opinion. *Sch. & Soc.*, 1943, **57**, 333–335.
56. Morgan, J. J. B., & Morton, J. T. The distortion of syllogistic reasoning produced by personal convictions. *J. soc. Psychol.*, 1944, **20**, 39–59.
57. Morgan, J. J. B. Attitudes of students toward the Japanese. *J. soc. Psychol.*, 1945, **21**, 219–246.
58. Murray, H. A. *Explorations in personality.* New York: Oxford Univ. Press, 1938.
59. Murray, H. A., & Morgan, C. D. A clinical study of sentiments. I and II. *Genet. Psychol. Monogr.*, 1945, **32**, 3–311.
60. Murphy, G., & Likert, R. *Public opinion and the individual.* New York: Harpers, 1937.
61. Newcomb, T. M. *Personality and social change,* New York: Dryden Press, 1943.
62. Newcomb, T. M. The influence of attitude climate upon some determinants of information. *J. abnorm. soc. Psychol.*, 1946, **41**, 291–302.
63. Noland, E. W. Factors associated with absenteeism in a south central New York State industry. Ph.D. thesis on file in the Cornell University Library, 1944.
64. Parrish, J. A. The direct and indirect assessment of attitudes as influenced by propagandized radio transcriptions. Master's thesis, The Ohio State University, 1948.
65. Proshansky, H. A projective method for the study of attitudes. *J. abnorm. soc. Psychol.*, 1943, **38**, 383–395.
66. Robbins, I. Point of view and quality of thought in attitude measurement. *Improving Educational Research*, pp. 52–56. American Educ. Res. Ass. 1948 Official Report, Washington, D. C.
67. Rose, Arnold. *Studies in reduction of prejudice.* Chicago: American Council on Race Relations, 1948.
68. Rotter, J. B., & Willerman, B. The incomplete sentences tests as a method of studying personality. *J. consult. Psychol.*, 1947, **11**, 43–48.
69. Ruch, F. L. *Psychology and life.* (1st Ed.). Chicago: Scott, Foresman and Company, 1937.
70. Saadi, M., & Farnsworth, P. R. The degrees of acceptance of dogmatic statements and preferences for their

supposed makers. *J. abnorm. soc. Psychol.*, 1934, **29**, 143–150.

71. Seeleman, V. The influence of attitude upon the remembering of pictorial material. *Arch. Psychol., N. Y.*, 1940–1941, **36**, No. 258.

72. Seeman, M. Moral judgments: a study in racial frames of references, *Amer. sociol. Rev.*, 1947, **12**, 404–411.

73. Sherif, M. A study of some social factors in perception. *Arch. Psychol., N. Y.*, 1935, No. 187.

74. Sherif, M., & Cantril, H. *The psychology of ego-involvements.* New York: Wiley, 1947.

75. Sletto, R. F. *Construction of personality scales by the criterion of internal consistency.* Hanover: The Sociological Press, 1937.

76. Smith, G. H. Beliefs in statements labeled fact and rumor. *J. abnorm. soc. Psychol.*, 1947, **42**, 80–90.

77. Sweet, L. *The measurement of personal attitudes in younger boys.* New York: The Association Press, 1929.

78. Thurstone, L. L., & Chave, E. J. *The measurement of attitude.* Chicago: Univ. Chicago Press, 1929.

79. Travers, R. M. W. A study in judging the opinions of groups. *Arch. Psychol., N. Y.*, No. 266, 1941.

80. Wallen, R. Individuals' estimates of group opinion. *J. soc. Psychol.*, 1943, **17**, 269–274.

81. Watson, G. B. The measurement of fair-mindedness. *Teach. Coll. Contr. Educ.* No. 176. New York: Teachers Coll., Columbia Univ., 1925.

82. Williams, Robin M., Jr. *The reduction of intergroup tensions.* New York: Social Science Research Council Bull., No. 57, 1947.

83. Wolff, H. A., Smith, C. E., & Murray, H. A. The psychology of humor; 1. a study of race disparagement jokes. *J. abnorm. soc. Psychol.*, 1934, **28**, 341–365.

84. Zeligs, R. Racial attitudes of children. Sociol. soc. Res., 1937, **21**, 361–371.

Received May 1, 1949.

AN EXPERIMENTAL ANALOGUE OF REPRESSION
I. HISTORICAL SUMMARY[1]

ANCHARD FREDERIC ZELLER
The Johns Hopkins University

INTRODUCTION

At the heart of the theoretical psychoanalytic structure is the concept of repression, which is the mechanism by which painful or unpleasant material is excluded from consciousness and motor expression. In his earlier writings Freud (**17**) stressed the importance of repression as the primary ego defense but in his later works he revised his concept by pointing out that repression is only one of several mechanisms which the ego may utilize in an effort to avoid unpleasantness. The ego may indulge in flight, it may resort to condemnation, in which case the material remains conscious, or it may resort to repression.

Repression is of two kinds: first, primal or archaic repression which denies entrance into consciousness of some archaic idea attached to instinctual strivings which are unacceptable to the ego, and second, after-expulsion or repression proper which pushes from consciousness material which, although once conscious, has in some way become associated with primally repressed material. Repression is not forgetting. It is an active process which requires the exertion of constant energy by the ego. This is true for both primal repression and after-expulsion. Repression may be complete or partial. If it is complete there is very little hope of investigating it experimentally. If it is partial the repressed material may remain as it was, ready to reappear in consciousness if the repression is removed, or the force attached to it may manifest itself in consciousness in the form of neurotic symptoms. Incomplete repression lends itself more readily to experimental investigation. From the standpoint of the clinician, after-explusion is of more importance than primal repression, since most observed cases of repression are examples of incomplete after-explusion. That is, an individual is unable to recall material which he has at one time known but which has in some way become unacceptable, but the memory for which may be restored by appropriate treatment. Although seldom explicitly stated, this is the concept of repression which has served as a basis for most of the experiments in the field. A great many experiments have been designed, first, to demonstrate the phenomenon and second, to quantify the factors involved. Such a demonstration must of necessity fulfill three requirements. First, it must demonstrate that the material in question has been learned by the individual. Second, it must demonstrate that the introduction of an inhibiting factor causes inability to recall or a significant decrease in the recall of the material. Third, it must show that the removal of the inhibiting factor results in the reinstatement of the ability to recall the material. As will be shown in this review of the literature there has as yet been no clear cut laboratory demonstration which includes all of these steps.

---

[1] The present paper is a portion of a dissertation presented to the Faculty of Philosophy of the School of Higher Studies, The Johns Hopkins University, in partial fulfillment of the requirements for the Ph.D. degree. The writer is indebted to Professors Stanley B. Williams and W. C. H. Prentice for advice and assistance during the course of the study.

EXPERIMENTAL STUDIES OF REPRESSION

Questionnaire method. The first experimental attack was made by F. W. Colgrave (**10**) in 1898 when he administered a questionnaire to a group of school children which contained the question, "Do you recall pleasant or unpleasant experiences better?" He concluded from his results that pleasant items were better recalled than unpleasant. As critics have pointed out, the questionnaire method is not a highly valid measure, but a start had been made. Kowalewski (**30**) and again Susukita (**81**) as late as 1935 also utilized the questionnaire in the study of this problem.

Associations with sensory stimuli. In 1905 Gordon (**20**), working in Külpe's laboratory, initiated another type of attack on the problem by associating sensory stimuli with material to be learned and later testing recall. She used colored figures paired with protocols, and administered a recognition test after three weeks. She found no difference between pleasant, unpleasant, and indifferent items. Tait (**82**), using color recognition, found a slight superiority for memory associated with pleasant colors. Anderson and Bolton (**1**) tested the recognition and recall of odors with nonsense names attached and found little difference between pleasant, unpleasant, and indifferent items. Gordon (**21**) again attacked the problem in 1925, this time using pleasant and unpleasant odors followed by a memory test. Again she found no difference between memory for pleasant and unpleasant items. Kenneth (**28**) used odors as a free association stimulus. He found more pleasant than unpleasant associations, but the results are not especially relevant to repression. Frank and Ludvigh (**16**) paired odors with nonsense syllables and found that pleasant associations were recalled oftener. The most complete experiment on memory value of pleasant and unpleasant items was conducted by Ratliff (**55**) using sensory stimulation paired with numbers. Three types of sensory stimuli, odors, colors, and consonant chords, were paired with the numbers. He found that with visual and auditory stimulation the pleasant was recalled more often than the unpleasant, but for olfaction the unpleasant recall was superior.

Recall of experiences. Another approach extensively used to study the relation of affect to memory was initiated by Kowalewski (**31**) in 1908 when, the day following a Christmas vacation, he instructed his students, "Write down whatever pleased you yesterday." His results indicated more pleasant than unpleasant experiences. A recall ten days later gave similar results. He interpreted his results to mean that the pleasant is better retained than the unpleasant. As critics were quick to point out, however, he assumed that pleasant and unpleasant experiences were equal in number, which assumption has more recently been proven erroneous. Other investigators employing the recall of holiday experiences have been Wohlgemuth (**91**), Meltzer (**40, 41, 43**), Jersild (**27**), Cason (**8**), Waters and Leeper (**87**), Menzies (**44**), O'Kelly and Steckle (**48**), and Steckle (**79**). The studies of Wohlgemuth (**91**) and Cason (**8**), two of the best controlled, indicate no difference between recall of pleasant and unpleasant associations when the fact that there are actually more pleasant than unpleasant experiences is taken into account, while the results of the other investigations indicate only a slight tendency for greater recall of pleasant experiences.

Kowalewski (**30, 31**) concluded from his results that the recall of affective material is related to personality types. He interpreted his data to mean that optimists remember proportionally more pleasant experiences while pessimists tend to remember proportionally more unpleasant experiences. This possibility has been investigated by Meltzer (**40, 43**), O'Kelly and Steckle (**48**), and Steckle (**79**) with inconclusive results. Flugel in 1917 (see Flugel, 1925, **14**) attempted to establish an "Algedonic Ratio" of pleasant to unpleasant for the individual. This was again attempted by Meltzer (**41**) in 1930.

Other experiments which have used the recall of past experience have been conducted by Gordon (**22**) and Thompson (**83**) who used the recall of childhood experiences. Gordon found no evidence for a greater percentage of pleasant recall, but Thompson found evidence for a pleasant-unpleasant differential in favor of the pleasant.

Discrete associations. In addition to the experiments in which preceding events have been recalled and re-recalled, a number of experiments have utilized other methods to test memory for affectively toned experience. Peters (**49**) had subjects answer to a stimulus word with an experience. He found 65 per cent of total recalls were pleasant in nature. Peters and Nemecek (**50**) used the same technique with comparable results. Griffitts (**24**) asked for responses to stimulus words and found a difference in favor of pleasant associations. Stagner (**77**) in a well-controlled experiment in which a single pleasant and unpleasant event were recorded, together with their associations, found what he called evidence for an active repression when the events were again presented to the subjects with instructions to recall the associations.

Reaction time. The effect on reaction time of various affectively toned material has been investigated by Birnbaum (**5**), Tolman (**84**), Tolman and Johnson (**85**), Baxter, Yamada and Washburn (**4**), Morgan, Mull and Washburn (**46**), and Smith (**76**). All but Birnbaum found positive results, viz., pleasant tone decreases reaction time.

Memory for word lists. The number of experiments using this method has been large. Chaney and Lauer (**9**), Lynch (**36**), Cason (**8**), Stagner (**78**), Balken (**2**), Bunch and Wientge (**6**), Silverman and Cason (**75**), White and Ratliff (**90**), Carter (**7**), White (**89**), Gilbert (**18**), Lanier (**33**), and Pintner and Forlano (**51**) have all worked with this type of material, using periods of recall up to two years. None of these experiments has given a conclusive answer to the problem. They have all tended to show, however, that the problem of affect and recall is not a simple one, but rather a very complex phenomenon depending on many factors, such as sex, age, social status, intelligence, etc. (see Gilbert, 1938).

Miscellaneous studies. A few other studies have been conducted which do not readily fall into the above categories. Laird (**32**) asked his subjects to write the names of ten friends and check those liked and disliked. Later he asked the subjects to list ten names as fast as possible. A comparison of the two lists indicated that some subjects tended to list more disliked persons in the second list than did others. He interpreted his results in terms of optimistic and pessimistic personality types but found no evidence for repression. Fox (**15**) had his subjects learn pleasant and unpleasant sonnets and found those preferred were better retained. Stone (**80**) used photographs to which names liked or disliked by the subjects were attached. He found that those associated with pleasant names were better recalled. McMullin (**39**) found evidence that interpolated activity of an affective nature could produce differences in the amount of retroactive inhibition.

Review articles or discussions of various aspects of pleasant-unpleasant experiments have been presented by Meltzer (**42**), Moore (**45**), Sears (**64, 67, 68, 69**), Gilbert (**19**), McGeoch (**37**), Rapaport (**54**), and Prentice (**53**). As can be seen from the experiments reviewed so far, very little has been contributed to an understanding of the problem of repression as outlined by Freud. Of the experiments mentioned so far, the score stands 32 to 14 in favor of more effective recall of pleasant than of unpleasant experiences, with 5 experiments having neutral results. These results, however, can not be taken as an indication of repression.

As Meltzer (**42**) has pointed out, most of the experiments before 1930 were conducted on the false assumption that the pleasant and unpleasant events were of equal frequency. Henderson (**25**) as early as

1911, however, recognized the fallacy and interpreted this as failing to support the idea that the unpleasant was repressed. Myers (**47**) also argued that we do not forget the disagreeable, but that there are more agreeable than disagreeable experiences in life. Flugel (**14**) also recognized this as did Griffitts (**24**) and Wohlgemuth (**91**). In none of the studies has any attempt been made to restore recall when repression has been assumed as basic to the pleasant-unpleasant differential. In most cases repression has been equated to differential forgetting, whereas it could have been more reasonably interpreted as differential learning. In most cases repression has not been conceived as the dynamic process outlined by Freud, and in no case has an adequate test been designed.

Frame of reference. There have, however, been a number of studies which have recognized the importance of learning set. Following the work of Bartlett (**3**) a number of studies have been conducted in which controversial material about which the subjects were known to have specific opinions was presented and then recalled later. As early as 1928 Zillig (**93**) gave both men and women a number of selections to read, the content of some being favorable and others being unfavorable to women. The women recalled a much greater percentage of favorable items about themselves than the men did about the women. Studies following the same plan have been conducted by Watson and Hartmann (**88**) using theistic and atheistic material, Seeleman (**70**) using attitude toward the negro, Edwards (**11**) with pro- and anti- New Deal material, Levine and Murphy (**34**), pro- and anti-communistic literature and Postman and Murphy (**52**) using United Nations vs. Axis power attitudes. Wallen (**86**) used his subjects' own opinion of themselves, compared with bogus group opinions, and found they remembered the items which agreed with their own opinions better than those which did not. Shaw (**73**) and Shaw and Spooner (**74**) repeated the same experiment with modifications and obtained comparable results. All these studies indicate that attitude and preconceptions influence memory. Edwards (**12**) has presented an argument in favor of a "frame of reference" determinant in forgetting, regardless of whether the material is intrinsically pleasureful or otherwise. Steckle (**79**) presents a similar argument. Again, however, it should be noted that the analysis is of differential forgetting, not of an active, removable repression.

Specific studies. There have been a number of experiments based on different approaches to the problem which have shown considerably more insight into the nature of the problem and more imagination in the design of the experiment. Koch (**29**) using a number of psychology students had them rate their satisfaction with grades on quizzes. Five weeks later a recall test of the grades indicated the grades with the most affect, either positive or negative, were better recalled. In 1936 Sears (**64**) presented an excellent review of functional abnormalities of memory in which he pointed out that none of the experiments had fulfilled the conditions of a true test of repression, since the fundamental assumption of the experiments had been that pleasantness and unpleasantness of an intellectual or sensory nature is equivalent to unpleasantness in terms of ego threat. In 1937 Sears (66) presented a study in which subjects were given two tasks, learning nonsense syllables and sorting cards. A list of syllables was learned, followed by success or failure at card sorting, and this was in turn followed by the learning of a second list. The second list learned by the successful card sorters was significantly better than the learning of the same list by those who failed at card sorting. Sears interprets this as evidence of repression. It might be pointed out that the difference could well be attributed to lesser motivation rather than to an active repression. No attempt was made to remove the repression.

One of the better demonstrations of repression has been given by Huston, Shakow, and Erickson (26) using the technique developed by Luria (35) in which the "associations of higher central nervous system processes are associated with voluntary movement so that the conflicts in the former are disclosed in the latter." The subjects were hypnotized and were told that they had participated in some event in a manner out of keeping with their normal standard of ethics. Stimulus words, some neutral and some related to the suggested experience, were read to the subjects who had a post-hypnotic amnesia for the suggestions. There were significant differences in reaction to the words associated with the hypnotic suggestions. The subjects were then rehypnotized and the suggestion removed. Retests indicated the effects had disappeared. This experiment comes close to a demonstration of repression, but the use of hypnosis as a repression medium makes it difficult to interpret.

Flanagan (13) and Sharp (71) at the University of Chicago used sexual and profane word pairs, respectively, and found significant differences in the learning and recall of the two types of material. It should be pointed out in this connection that no attempt was made to equate the affective and control material for difficulty in learning, and the experimental situation was such that embarrassment over the nature of the material to be reproduced undoubtedly contributed a great deal to the results. Sharp (72) in 1938, using neurotic subjects and material taken from case histories, found curves of retention which would indicate that pleasant material is more readily recalled after a time while unpleasant material is repressed (forgotten).

McGranahan (38) in an interesting interpretation of the Freudian concept, points out that repression is a process which can be carried out more efficiently at a conscious level by a well-integrated personality. To demonstrate this, he forbade his subjects to name a color on penalty of being shocked. The subject then answered to a series of stimulus words to which color responses are common. The well-integrated individuals "repress" the color responses while those who fear the shock make more color responses.

In 1930 Meltzer (**42**) concluded after a critical review of the studies on feeling and memory that the difficulty with repression studies was that the investigators were looking for universal responses. Rosenzweig (**56, 57, 58**), and Rosenzweig and Mason (**61**) pointed out that the lack of acceptable results in previous studies was due to a lack of clear understanding of the problem. They further point out the fallacy of assuming that sensory pleasantness and unpleasantness is the same as conative pleasantness and unpleasantness. Further, repression operates on a conative type of material which is negative in hedonic tone, the negative affect being determined by the conflict of the material with ego-supporting drives, such as self-respect, etc. They cite the experiments of Zeigarnik (**92**) as a prototype of the repression experiment and set up a series of investigations in which children were allowed to complete one half of the tasks and fail on one half. Recall was better for the completed tasks. Varying the task set, Rosenzweig (**59, 60**) found that ego-involvement led to recall of completed tasks, while task orientation led to better recall of uncompleted tasks. Two experiments conducted by Sarason and Rosenzweig (**62, 63**) in connection with Rosenzweig's triadic hypothesis further attempted to correlate repression with hypnotizability and impunitive reactions to conflict. Rosenzweig interprets his evidence as lending support to his hypothesis.

Gould (**23**) designed an experiment which came close to a test of repression. The subject was required to choose one of two tasks. The choice supposedly revealed a good or bad character trait of the subject, the character traits "revealed" having been predetermined by the investigator regardless of the task chosen. The subject was then told what the choice revealed about him and was allowed to complete the task, then make another choice until six had been chosen. The subject was then asked to list all of the tasks. Following this, the nature of the experiment was explained to the subject. As an incidental remark, Gould notes that after the explanation some of the subjects remembered more of the tasks than before.

SUMMARY AND CONCLUSIONS

A review of the literature shows that there has been no experiment which has fulfilled the criterion of a laboratory test of repression. Most experiments have been directed toward testing the relationship between affect and recall. The majority of experiments have been found to fall in eight categories:

1. Questionnaire method: this was the earliest used method and is unsuited to the problem.
2. Associations with sensory stimuli: the chief criticism of this method

was that it assumed the equivalence of sensory unpleasantness with ego unpleasantness.

3. Recall of experiences: many of these experiments have assumed actual numbers of pleasant and unpleasant experiences to be equal.

4. Discrete associations: association tests have in general shown more pleasant associations.

5. Reaction time: these experiments have shown shorter reaction time with pleasant material.

6. Memory for word lists: this method, the most frequently used, has yielded nothing significant.

7. Frame of reference: these studies have shown a selective effect of attitude on memory.

8. Specific studies: these studies, specifically reviewed in the article, while showing a clearer conception of the problem, have failed to demonstrate repression.

All reviews and criticisms of the problem have agreed that no adequate test of repression has been made. In most studies the repression has been conceived as differential forgetting. The most significant criticism has not been specifically formulated, namely, that *no test of repression can be considered adequate until the removal of the repression factor has resulted in the restoration to consciousness of the repressed material.* Any experiment which does not include this crucial step is not complete and the results can be attributed to other factors such as set, differential learning, differential motivation, practice, etc. rather than to active repression.

A proposed experimental design to test the hypothesis that repression proper is a process which inhibits recall without, however, destroying the memory for the subject matter, and that the memory can be restored by the elimination of the repression factor is presented below.

		Control		*Experimental*
I.	Step 1.	Learning	=	Learning
	Step 2.	Retention test	=	Retention test
		Time interval		
II.	Step 3.	Retention test	=	Retention test
	Step 4.	Neutral task		Repression
	Step 5.	Retention test	>	Retention test
		Time interval		
III.	Step 6.	Retention test	>	Retention test
	Step 7.	Neutral task		Removal of repression
	Step 8.	Retention test	=	Retention test

If the hypothesis is supported, the results indicated in the design can be expected. Any other results, although not disproving repression, would cast real doubt on the validity of the concept.

The design involves two equated groups, a control and an experimental group. The individuals in each group are exposed to some material (Step 1) which they are to learn. Their degree of mastery is then measured with a retention test (Step 2). This satisfies the first requirement of an experimental test of repression proper, namely that the material be learned.

After a time interval another test of retention of the originally learned material is administered to the individuals in both groups (Step 3). Up to this point the performance of the groups should show no significant difference. At this point the procedure varies. The individuals in the experimental group are subjected to an ego threat associated with some activity while the individuals in the control group perform the same activity with no induced affect (Step 4). After this another retention test of the originally learned material is administered (Step 5). If the average retention in Step 5 is significantly better for the control group, the second criterion of an adequate experimental test of repression has been met.

After another time interval the subjects are again asked for a measure of retention of the originally learned material. (Step 6). This is followed by the same activity associated with the ego threat in Step 4, only this time there is no ego threat for either the experimental or the control group. Restoration of memory following the change in ego value of the task would constitute the third and crucial step in the experimental demonstration of repression proper.

BIBLIOGRAPHY

1. Anderson, A. C., & Bolton, F. J. The inhibition of the unpleasant. *J. abnorm. soc. Psychol.*, 1925, **20**, 300–302.
2. Balken, E. R. Affective, volitional and galvanic factors in learning. *J. exp. Psychol.*, 1933, **16**, 115–128.
3. Bartlett, F. C. *Remembering: a study in experimental and social psychology.* New York: Macmillan; Cambridge, Eng.: Univ. Press, 1932.
4. Baxter, M. F., Yamada, K., & Washburn, M. F. Directed recall of pleasant and unpleasant experiences. *Amer. J. Pyschol.*, 1919, **30**, 300–302.
5. Birnbaum, K. Ueber den Einfluss von Gefühlsfaktoren auf die Assoziationer. *Monatsch f. Psychiat. u. Neurol.*, 1912, **32**, 95–123.
6. Bunch, M. E., & Wientge, E. The relative susceptibility of pleasant, unpleasant, and indifferent material to retroactive inhibition. *J. exp. Psychol.*, 1933, **9**, 157–178.
7. Carter, H. D. Effect of emotional factors upon recall. *J. Psychol.*, 1936, **1**, 48–55.
8. Cason, H. The learning and retention of pleasant and unpleasant activities. *Arch. Psychol., N. Y.*, 1932, No. 134. Pp. 96.
9. Chaney, Ruth, & Lauer, A. R. The influence of affective tone on learning and retention. *J. educ. Psychol.*, 1929, **20**, 287–291.

10. Colgrave, F. W. Individual memories. *Amer. J. Psychol.*, 1898–99, **10**, 228–255.
11. Edwards, A. L. Political frames of reference as a factor influencing recognition. *J. abnorm. soc. Psychol.*, 1941, **36**, 34–61.
12. Edwards, A. L. The retention of affective experiences—a criticism and restatement of the problem. *Psychol. Rev.*, 1942, **49**, 43–53.
13. Flanagan, D. E. The influence of emotional inhibition on learning and recall. Unpublished Master's thesis, Univ. Chicago, 1930.
14. Flugel, J. C. A quantitative study of feeling and emotion in everyday life. *Brit. J. Psychol.*, 1925, **15**, 318–355.
15. Fox, C. The influence of subjective preference on memory. *Brit. J. Psychol.*, 1922–23, **13**, 398–405.
16. Frank, J. P., & Ludvigh, E. J. The retroactive effect of pleasant and unpleasant odors on learning. *Amer. J. Psychol.*, 1931, **43**, 102–108.
17. Freud, S. Repression. In *Collected papers*, Vol. 4, 84–97. London: Hogarth Press, 1925.
18. Gilbert, G. M. The age difference in the hedonistic tendency in memory. *J. exp. Psychol.*, 1937, **21**, 433–441.
19. Gilbert, G. M. The new status of experimental psychology on the relationship of feeling to memory. *Psychol. Bull.*, 1938, **35**, 26–35.
20. Gordon, Kate. Ueber das Gedachtnis fur Affective bestimmte Eindrucke. *Arch. ges. Psychol.*, 1905, **4**, 437–458.
21. Gordon, Kate. Recollection of pleasant and unpleasant odors. *J. exp. Psychol.*, 1925, **8**, 225–239.
22. Gordon, Kate. A study of early memories. *J. Delinqu.*, 1928, **12**, 127–132.
23. Gould, R. Repression experimentally analyzed. *Character and Pers.*, 1942, **10**, 259–288.
24. Griffitts, C. H. Results of some experiments on affection, distributions of associations and recall. *J. exp. Psychol.*, 1920, **3**, 447–464.
25. Henderson, E. N. Do we forget the disagreeable? *J. Phil. Psychol. sci. Meth.*, 1911, **8**, 432–438.
26. Huston, P. E., Shakow, D., & Erickson, M. H. A study of hypnotically induced complexes by means of the Luria technique. *J. gen. Psychol.*, 1934, **11**, 65–97.
27. Jersild, A. Memory for the pleasant as compared with the unpleasant. *J. exp. Psychol.*, 1931, **14**, 284–288.
28. Kenneth, J. H. An experimental study of affects and associations due to certain odors. *Psychol. Monog.*, 1927, No. 37. Pp. 64.
29. Koch, H. L. The influence of some affective factors upon recall. *J. gen. Psychol.*, 1930, **4**, 171–190.
30. Kowalewski, A. *Studien zur Psychologie des Pessimismus.* Wiesbaden, 1904.
31. Kowalewski, A. *Schopenhauer und sein Weltanschauung.* Berlin, 1908.
32. Laird, D. The influence of likes and dislikes on memory as related to personality. *J. exp. Psychol.*, 1923, **6**, 294–300.
33. Lanier, L. H. Memory for words differing in affective value. *Psychol. Bull.*, 1940, **37**, 492–493. (Abstract.)
34. Levine, J. M., & Murphy, G. The learning and forgetting of controversial material. *J. abnorm. soc. Psychol.*, 1943, **38**, 507–517.
35. Luria, A. R. *The nature of human conflicts.* (Trans. by W. H. Gantt.) New York: Liveright, 1932.
36. Lynch, C. A. The memory value of certain alleged emotionally toned words. *J. exp. Psychol.*, 1932, **15**, 298–315.
37. McGeoch, J. O. *The psychology of human learning.* New York: Longmans, Green, 1946.
38. McGranahan, D. V. A critical and experimental study of repression. *J.*

abnorm. soc. Psychol., 1940, **35**, 212–225.

39. McMullin, T. E. A study of the affective nature of the interpolated activity as a factor in producing different relative amounts of retroactive inhibition in recall and recognition. *J. exp. Psychol.*, 1942, **30**, 201–215.
40. Meltzer, H. The forgetting of pleasant and unpleasant experiences in relation to intelligence and achievement. *J. soc. Psychol.*, 1930, **2**, 217–227.
41. Meltzer, H. Individual differences in forgetting pleasant and unpleasant experience. *J. educ. Psychol.*, 1930, **21**, 399–409.
42. Meltzer, H. The present status of experimental studies of the relation of feeling to memory. *Psychol. Rev.*, 1930, **37**, 124–139.
43. Meltzer, H. Sex differences in forgetting pleasant and unpleasant experiences. *J. abnorm. soc. Psychol.*, 1931, **25**, 450–464.
44. Menzies, R. The comparative memory value of pleasant, unpleasant and indifferent experiences. *J. exp. Psychol.*, 1936, **18**, 267–279.
45. Moore, E. H. A note on the recall of the pleasant vs. the unpleasant. *Psychol. Rev.*, 1935, **42**, 214–215.
46. Morgan, E., Mull, H. K., & Washburn, M. F. An attempt to test moods or temperaments of cheerfulness and depression by directed recall of emotionally toned experiences. *Amer. J. Psychol.*, 1919, **30**, 302–304.
47. Myers, G. C. Affective factors in recall. *J. Phil. Psychol. sci. Meth.*, 1915, **12**, 85–92.
48. O'Kelly, L. I., & Steckle, L. C. The forgetting of pleasant and unpleasant experiences. *Amer. J. Psychol.*, 1940, **53**, 432–434.
49. Peters, W. Gefühl und Erinnerung; Beiträge zur Erinnerungsanalyse. *Psychol. Arbeit.*, 1911, **6**, 197–260.
50. Peters, W., & Nemecek, O. Massenversuche über Erinnerungassoziationen. *Fortschr. Psychol. Anwend.*, 1914, **2**, 226–245.
51. Pintner, R., & Forlano, G. The influence of pleasantly and unpleasantly toned words on retention. *J. soc. Psychol.*, 1940, **11**, 147–148.
52. Postman, L., & Murphy, G. The factor of attitude in associative memory. *J. exp. Psychol.*, 1943, **33**, 228–238.
53. Prentice, W. C. H. The interruption of tasks. *Psychol. Rev.*, 1944, **51**, 329–340.
54. Rapaport, D. *Emotions and memory.* Baltimore: Williams and Wilkins, 1942.
55. Ratliff, M. M. The varying function of affectively toned olfactory, visual and auditory cues in recall. *Amer. J. Psychol.*, 1938, **51**, 695–699.
56. Rosenzweig, S. The recall of finished and unfinished tasks as affected by the purpose with which they were performed. *Psychol. Bull.*, 1933, **30**, 698. (Abstract.)
57. Rosenzweig, S. The experimental study of repression. In H. Murray (Ed.), *Explorations in personality.* New York: Oxford Univ. Press, 1938. Pp. 472–491.
58. Rosenzweig, S. The experimental measurement of types of reaction to frustration. In H. Murray (Ed.), *Explorations in personality.* New York: Oxford Univ. Press, 1938. Pp. 585–599.
59. Rosenzweig, S. Need-persistive and ego-defensive reactions to frustration as demonstrated by an experiment on repression. *Psychol. Rev.*, 1941, **48**, 347–349.
60. Rosenzweig, S. An experimental study of "repression" with special reference to need-persistive and ego-defensive reactions to frustration. *J. exp. Psychol.*, 1943, **32**, 64–74.

61. Rosenzweig, S., & Mason, G. An experimental study of memory in relation to the theory of repression. *Brit. J. Psychol.*, 1934, **24**, 247–265.

62. Rosenzweig, S., & Sarason, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. I. Correlational approach. *Character and Pers.*, 1942, **11**, 1–19.

63. Sarason, S., & Rosenzweig, S. An experimental study of the triadic hypothesis: reaction to frustration, ego-defense, and hypnotizability. II. Thematic Apperception Approach. *Character and Pers.*, 1942, **11**, 150–165.

64. Sears, R. R. Functional abnormalities of memory with special reference to amnesia. *Psychol. Bull.*, 1936, **33**, 229–274.

65. Sears, R. R. An experimental test of one phase of the hypothecated repression sequence. *Psychol. Bull.*, 1936, **33**, 744. (Abstract.)

66. Sears, R. R. Initiation of the repression sequence by experienced failure. *J. exp. Psychol.*, 1937, **20**, 570–580.

67. Sears, R. R. Non-aggressive reactions to frustration. *Psychol. Rev.*, 1941, **48**, 343–348.

68. Sears, R. R. Survey of objective studies of psychoanalytic concepts. *Soc. Sci. Res. Coun. Bull.* No. 51, 1943.

69. Sears, R. R. Experimental analysis of psychoanalytic phenomena. In J. McV. Hunt (Ed.), *Personality and the behavior disorders.* New York: Ronald Press, 1944. Pp. 306–332.

70. Seeleman, V. The influence of attitude upon the remembering of pictorial material. *Arch. Psychol., N. Y.*, 1940, **36**, No. 258.

71. Sharp, Agnes A. The influence of certain emotional inhibitions on learning and recall. Unpublished Master's thesis, Univ. Chicago, 1930.

72. Sharp, Agnes A. An experimental test of Freud's doctrine of the relation of hedonic tone to memory revival. *J. exp. Psychol.*, 1938, **22**, 395–418.

73. Shaw, F. J. Two determinants of selective forgetting. *J. abnorm. soc. Psychol.*, 1944, **39**, 434–445.

74. Shaw, F. J., & Spooner, A. Selective forgetting when the subject is not ego involved. *J. exp. Psychol.*, 1945, **35**, 242–247.

75. Silverman, A., & Cason, H. Incidental memory for pleasant, unpleasant and indifferent words. *Amer. J. Psychol.*, 1934, **46**, 315–320.

76. Smith, W. W. Experiments on memory and affective tone. *Brit. J. Psychol.*, 1921, **11**, 236–250.

77. Stagner, R. The redintegration of pleasant and unpleasant experiences. *Amer. J. Psychol.*, 1931, **43**, 463–468.

78. Stagner, R. Factors influencing the memory value of words in a series. *J. exp. Psychol.*, 1933, **16**, 129–137.

79. Steckle, L. C. Again—affect and recall. *J. soc. Psychol.*, 1945, **22**, 103–105.

80. Stone, A. R. The reaction of memory to affective states. *Amer. J. Psychol.* 1925, **36**, 112–123.

81. Susukita, T. Ueber das Gedächtnis für lust- und unlustbetonte Erlebniss im Alltagsleben. *Tohoku Psychol. Folia*, 1935, **3**, 187–204.

82. Tait, W. D. Effects of psychophysical attitudes on memory. *J. abnorm. soc. Psychol.*, 1913–14, **8**, 10–38.

83. Thompson, R. H. An experimental study of memory as influenced by feeling tone. *J. exp. Psychol.*, 1930, **13**, 462–467.

84. Tolman, E. C. Retroactive inhibition as affected by conditions of learning. *Psychol. Monogr.*, 1918, No. 107, 187–195.

85. Tolman, E. C., & Johnson, I. A note on association time and feeling.

Amer. J. Psychol., 1918, **29**, 187–195.

86. Wallen, D. Ego involvement as a determinant of selective forgetting. *J. abnorm. soc. Psychol.*, 1942, **37**, 20–29.
87. Waters, R. H., & Leeper, R. The relation of affective tone to the retention of experiences in everyday life. *J. exp. Psychol.*, 1936, **19**, 203–215.
88. Watson, W. S., & Hartmann, G. W. The rigidity of a basic attitudinal frame. *J. abnorm. soc. Psychol.*, 1939, **34**, 314–335.
89. White, M. M. Some factors influencing recall of pleasant and unpleasant words. *Amer. J. Psychol.*, 1936, **48**, 134–139.
90. White, M. M., & Ratliff, M. M. The relation of affective tone to learning and recalling words. *Amer. J. Psychol.*, 1934, **46**, 92–98.
91. Wohlgemuth, A. The influence of feeling on memory. *Brit. J. Psychol.*, 1923, **13**, 405–416.
92. Zeigarnik, B. Über das Behalten von erledigten und unerledigten Handlungen. *Psychol. Forsch.*, 1927, **9**, 1–85.
93. Zillig, M. Einstellung und Aussage. *Z. Psychol.*, 1928, **106**, 58–106.

Received June 30, 1949.

ON THE CIRCULARITY OF THE LAW OF EFFECT

PAUL E. MEEHL
University of Minnesota

In his recent review on "The History and Present Status of the Law of Effect," Postman (**19**) lays considerable emphasis on the problem of "circularity" which he sees as crucial in the formulation of the law. He says:

> Whereas some critics were most concerned with the mechanisms mediating effect, others focussed their attention on the nature of the satisfiers and annoyers to which reference is made in Thorndike's law. Although Spencer and Bain, in whose tradition Thorndike continued, frankly invoked pleasure and pain as agents responsible for the fixation and elimination of responses, Thorndike's law has been a law of *effect*, not *affect*. He carefully defines satisfiers and annoyers in terms independent of subjective experience and report. "By a satisfying state of affairs is meant one which the animal does nothing to avoid, often doing such things as to attain and preserve it. By a discomforting state of affairs is meant one which the animal avoids and abandons." Although admittedly free of hedonism, such a definition of satisfiers and annoyers has faced another serious difficulty: the danger of circularity. The critic may easily reword the definition to read: "The animal does what it does because it does it, and it does not do what it does not do because it does not do it." This *reductio ad adsurdum* is probably not entirely fair, but it points up the danger of the definition in the absence of an *independent* determination of the nature of satisfiers and annoyers. The satisfying or annoying nature of a state of affairs can usually be determined fully only in the course of a learning experiment and cannot then be invoked as a causal condition of learning without circularity. In their experimental work Thorndike and his associates have made no significant attempts to establish the satisfying or annoying nature of their rewards and punishments independently of the learning experiment (**19**, p. 496).

And a little later Postman says:

> Stripped of virtually all defining properties and qualifications, the law does indeed have a very wide range of applicability but only at the expense of vagueness. The sum and substance of the argument now is that something happens in the organism (nervous system) after an act is performed. The fact that something happens influences further action. This something is, however, so little defined that it has almost no predictive efficiency. The O.K. reaction has no measurable properties, the conditions for its occurrence are so general as to embrace almost every conceivable situation. Hence the operation of O.K. reaction can be inferred only *ex post facto*, after learning has taken place. But here we are again impaled on the horns of the dilemma of circularity (**19**, p. 497).

And still further:

> In attempting to evaluate the controversy which has raged around the

> definition of satisfiers one is struck by the key importance of the hedonistic issue. Certainly hedonism is an immediate ancestor of the law, and now that the principle of effect has reached an uneasy maturity it is clear that it cannot deny its origin without sacrificing much of its vigor. When the law is stripped of hedonistic implications, when effect is not identified with tension-reduction or pleasure (as by Thorndike), the law of effect can do no more than claim that the state of affairs resulting from a response in some way influences future responses. Such a statement is a truism and hardly lends itself to the rigorous deduction of hypotheses and experimental tests. If a neohedonistic position is frankly assumed (as, e.g., by Mowrer) the law becomes an important tool for research, provided "satisfaction" is independently defined and not merely inferred from the fact that learning has occurred (**19**, p. 501).

Throughout Postman's paper this problem is constantly lurking behind the scenes even when the author does not single it out for specific mention. I am in complete agreement with Postman's final remark that "at the present state of our knowledge the law of effect as a monistic principle explaining all learning has not been substantiated," and Postman performs a service by emphasizing this problem of circularity in his discussion of the "law." I am inclined, however, to think that he has settled the question of circularity somewhat too easily, and that his settlement of it has an effect upon much of his argumentation. I gather from the above quotations that Postman looks upon any definition of effect or reinforcement in terms of the resulting change in response strength as "circular," where that word has a pejorative sense. If he is right in this it is very serious. While the law of effect has many difficulties, I do not believe that "circularity" is among them. To show this is the aim of the present paper.

I shall consider the problem of circularity in the law of effect as identical with the problem of circularity in the definition of *reinforcement* in instrumental conditioning. I take it that Postman does the same, since in the first quotation above he cites a passage from Hilgard and Marquis' *Conditioning and Learning*, where the two problems are considered together and with free interchange of the two terminologies. These authors say:

> It is apparent that no definition of effect provides an independent measure of the strength of reinforcement. The degree of satisfaction, of complacency, or of tension reduction has not been objectively determined. The strength of reinforcement can be given comprehensive definition only in terms of the amount of learning resulting from it. This is, of course, a circular definition, if strength of reinforcement is to be used as a factor determining degree of learning. A partial escape from circularity is achieved by the fact that a stimulus such as food which is found to be reinforcing in one situation will also be reinforcing in other situations, and with other animals (**9**, p. 83).

Bureau Edn'l. Psy. [illegible]

Writing in 1948, however, Hilgard states concerning Thorndike's "operational" definition of satisfiers and annoyers:

> These definitions are not circular, so far as the law of effect is concerned. That is, the states of affairs characterized as satisfying and annoying are specified independently of their influence upon modifiable connections. The law of effect then states what may be expected to happen to preceding modifiable connections which are followed by such specified states. The objection that Thorndike was lacking in objectivity in the statement of the law of effect is not a valid one (**8**, p. 24).

Hilgard is willing to let the concept of reinforcement (effect, satisfaction, reward) be introduced on the basis of behavior, but only because there are behavioral criteria of seeking and avoiding other than the effect of reinforcement upon *modifiable* connections. Whether this restriction is necessary needs to be considered carefully.

Skinner dismisses the whole problem in two sentences:

> A reinforcing stimulus is defined as such by its power to produce the resulting change. There is no circularity about this; some stimuli are found to produce the change, others not, and they are classified as reinforcing and nonreinforcing accordingly (**22**, p. 62).

Spence (**23**) takes essentially the same tack in his recent discussions of secondary reinforcement. The stimuli which impinge upon an organism may be divided, he says, into two classes: those which produce an increment in response strength, and those which do not. It seems from the several preceding quotations that there is a lack of agreement as to whether or not the law of effect or the principle of reinforcement involves an unavoidable circularity, or, if it does not, how circularity is to be avoided. In what follows, I make no claim to originality, since the essence of my development is contained in the previous quotations, together with the work of Tolman. But I feel it worthwhile to bring the arguments together in one context, and to show that the problem merits somewhat more extended treatment than is usually given it. Without claiming to present a definitive solution, I shall indicate the general direction which I believe the solution might take, and in the process introduce certain distinctions and terminological proposals which I feel might clarify our discussion and experimentation.

THE MEANING OF CIRCULARITY

It must be pointed out that there are two meanings of the word "circular" in common use. We have on the one hand circularity in *definition*, in which an unfamiliar term is defined by using other terms which are (directly or ultimately) defined by the term in question. There is no question of circularity in this sense in a definition of the

Skinner-Spence type. Let us accept as a crude preliminary formulation the following: "A reinforcing stimulus is one which increases the subsequent strength of responses which immediately precede it." The words *stimulus*,[1] *strength*, *increase* and *response* are all definable without any reference to the fact or theory of reinforcement. The definitions of these terms, particularly the term "response," present terrible difficulties; but I do not know of anyone who maintains that they involve the notion of reinforcement. Words such as these are current in the vocabulary of many kinds of psychological theorists who do not accept the Law of Effect as a principle of learning and in the absence of any indications to the contrary, I shall assume that we can tell what we mean by them. We can determine empirically when the strength of a response has increased without knowing anything about reinforcing stimuli, drives, satisfactions, and the like. It seems clear that the definition of a reinforcing stimulus in terms of its effect on response strength does not involve circularity in *this* sense.

The other meaning of the word circularity refers not to meanings (definition of terms) but to the establishment of propositions. We speak of *proofs* as being circular if it can be shown that in the process of establishing (proving) a proposition we have made use of the probandum. I am not aware that any responsible theorist has attempted to "prove" the Law of Effect in this way. It is true that those who look upon the law as fundamental are skeptical when they hear of a case of increase of response strength which does not *seem* to involve any obvious reinforcing consequences so that they begin to invent hypotheses to explain the results. There is no harm in this so long as the proposed explanations are in principle confirmable on the basis of some other experimental consequences, however remote. If an animal learns a response sequence without being given food, water, or any of the usual rewards, I suspect most Hullians would begin to talk about secondary reinforcement present in the situation. One can, of course, be careless with this kind of explanation, but there is nothing intrinsic to the concept that entails non-confirmability. The establishment of secondary reinforcing effects as explanations of a given experimental result consists in combining the facts known about primary reinforcers with facts about the animal's life history, in terms of which we understand how certain stimuli have acquired their secondary reinforcing powers. People on both sides of the present controversy over reinforcement theory are performing many different sorts of experiments in order to confirm or disconfirm the Law of Effect. It would seem that if the law of effect *were* being treated by anyone as a consequence of definition, or established by some hidden assumption of its truth, the experiments would not be going on.

---

[1] "Stimulus" will be used broadly to include "stimulus change," and stimulus configurations of all degrees of patterning and complexity.

CAN "REINFORCEMENT" BE INDEPENDENTLY DEFINED?

Nonetheless, when we think about this definition we feel uncomfortable. I do not think we have in mind either a circularity in definition or a begging-the-question fallacy, but some sort of peculiar pseudo-circularity in which it seems to us vaguely that the law *could* be "derived" from the proposed definition, even though no one in fact seems to be trying to do it this way. The problem can be stated very simply: How can we introduce the concept of reinforcement in terms of effect upon strength, and still have a "law of effect" or "principle of reinforcement" which has the empirical content that everybody seems to be taking for granted in experimentation?

1. Suppose we reject the Thorndike-Skinner-Spence procedure of defining reinforcement in terms of response strength, and decide to define the term quite independently of the learning process. The first possibility, which we shall dismiss rather dogmatically, is to do it subjectivistically in terms of pleasure, experiences of satisfaction, and the like. Aside from the general behavioristic objections, and the specific problems of measurement created, this approach is not feasible because it leaves us without any basis for speaking of reinforcing value in the case of that very important class of motivations that are unconscious or at least inadequately verbalized in the human case; and it makes impossible the establishment of reinforcing value in the case of lower organisms. At the present time there are probably very few psychologists who would consider this alternative seriously.

2. Secondly, we might try to define reinforcers in terms of certain physical properties on the stimulus side. I shall attempt to show below that this is a procedure which *follows* the introduction of the generic notion of a reinforcer, and which at a later stage becomes very important. But no one wants to group together an arbitrary class of physical objects or stimuli and call them "reinforcers," since the aim of our concept formation is to make possible the statement of laws. The possibility of identifying common physical properties of that large class of stimuli already grouped together as "rewarding" seems very remote. Besides, we would set up these properties or sets of properties by examining the members of the reinforcing class, which we already have set apart on some basis or other; and the question is: How have we arrived at the members of that class?

3. A third possibility, seen in the work of Hull, is to define reinforcement ultimately in terms of drive reduction, that is, in terms of the inner physiological events involved. Here again, I do not suppose that anyone would be able to give even the vaguest specification of the defining property of all neural events which are reinforcing. Even for the so-called primary physiological needs such as hunger, the evidence as to their exact physiological basis is most

incomplete. No psychologist today is willing to equate "hunger" with "stomach contractions," in the light of the experimentation on visceral denervations, specific sub-hungers, and the like. In other cases, we have practically no information on the neurophysiology, e.g., the neurophysiologic basis of the reinforcing effect of the presence of another organism, the turning off of a light in the Skinner box, or the going through of "exploratory" behavior on the other side of a grill. There is some reason to suppose that certain stimuli retain their secondary reinforcing value in the absence of the primary drive (2, 16), which complicates the problem further.

These considerations force a return to the *effect* of stimuli as a basis for specifying that they are reinforcers, and this leads to the paradox. If we define a reinforcing agent by its effect upon learning, then it seems that whenever learning is effected, we know ("by definition") that we have given a reinforcement. For surely, when the organism behaves, some stimulus change occurs, if nothing else than the proprioceptive effects of responding. If the behavior increases in strength, then these stimulus changes, which were in fact preceded by the response, are reinforcers. Hence, it seems that a definition of reinforcement in terms of an increase of habit strength makes the law tautological and devoid of factual content. This train of thought, which I am sure is familiar to most readers, seems obvious and straightforward. But I believe it can be shown to be mistaken, once the law is stated *explicitly* in the way we all really think of it *implicitly* when we perform experiments or try to explain a given case of learning.

AN EMPIRICAL DERIVATION OF REINFORCEMENT

Let us begin afresh by going to the behavior itself in a situation in which there is little or no disagreement as to what occurs. Consider a bright, inductively inclined Martian, who had never experienced any needs or satisfactions (except perhaps *n Cognizance*!) and who was observing the behavior of a rat in successive runnings in a T-maze. For the moment we shall simply consider a "standard rat," neglecting the individual differences in parameters and the accidents of personal histories that generate special secondary reinforcing properties. These refinements need to be added later, but as is usually the case will have to be added by being integrated into the whole structure of reinforcement theory, since we cannot treat everything at once. At the beginning, the Martian observes that the organism turns to the right or left with, let us say, about equal frequency. With further trials, a change occurs until finally the rat is responding close to 100% of the time by turning to the right. A Martian could obviously discover this with no notion of rewards, pleasure and the like. If he is ingenious enough to think of the possibility that the strength of a response might be influenced by the events that follow it in time, he would then proceed to investigate the changes that are contingent on this right turning.[2] He notes that when the rat turns to the right he brings about the following states of affairs on the stimulus side which he does not bring about when he turns to the left: He ends up nearer to the right-hand wall, which is painted green; he twists his own body to the right in responding; he ends up in a wooden box having knots in the wood; he ends up nearer the North pole; and to a dynamo on the other side of the campus; and he comes into the presence of a cup of sunflower seeds. These are the stimuli (stimulus changes) which are contingent on right turns. Is it possible that the gradual strengthening of the right turning is dependent upon one, some, or all of these changes following it? Our scientist from Mars would proceed to study a series of standard rats in the situation, altering the above variables systematically by usual inductive procedures. As a matter of empirical fact, he would discover that, within certain very wide limits, alterations in the first five have no effect. The sixth, the sunflower seeds, have a tremendous effect. He finds that he can alter the geographical direction, the direction of the body twist required, the wall color approached, etc.—that he can introduce all manner of modifications in the other factors; and so long as the sunflower seeds are presented, the rat will tend to go to where they are. On the other hand, if the sunflower seeds are omitted, and nothing else put in their place, a preference fails to develop as a function of these remaining differences.

But we have already greatly over-simplified. Actually, the Martian would discover that the effect of finding sunflower seeds in some cases is almost too slight to be detected; furthermore, even after a preference has been acquired, it may on some occasions fail to show itself. Now, it has already been apparent that when he comes upon these sunflower seeds, the rat behaves toward them in a characteristic way, that is, he ingests them. In seeking to understand the variability in the development and manifestation of a preference, one would notice a correlation between the strengthening of a preference and the rate, strength, and consistency of ingestive responses in the presence of the food. Identifying the same rat on successive days, it is found that on those days on which a preference already established broke down, very frequently the ingestive response in the presence of the sunflower seeds was at a very low or even zero strength. Failing to find anything varying in the maze situation itself to account for these differences, one can study the experiences of the animals between runs. Here appears a very striking correlate of both preference strength *and* the ingestive response in the maze: that which a human experimenter would call the "feeding schedule." The Martian would observe that when sunflower seeds were made available to the rats in their cages, they behave with respect to them in the same way as they do when they come upon the sunflower seeds in the goal box: namely, with ingestive responses. He would discover, again by systematic variation in these conditions, that such matters as the chemical nature of the substance made available, the periodicity of its availability, the lapse of time between when it was last available and the maze run; the rate of ingestion manifested at the moment of beginning deprivation (i.e., how close the rat was to satiety when interrupted), and so on, all exert an effect upon the maze response. By far the most intimate correlate would be the lapse of time since feeding. To quote Skinner again,

---

[2] Actually, no great ingenuity is involved here. Study of the events immediately *preceding* a run, e.g., the manner in which the experimenter handles the rat, what orientation he gives its head in placing it in the entry box, etc., would fail to reveal any systematic factor related to the direction of a preference. Considering this, together with the fact that before any runs have been made no preference exists, the Martian would be led to ask whether it is something that happens *after* the run (or during it) that affects the probability of a similar choice in subsequent runs.

> The problem of drive arises because much of the behavior of an organism shows an apparent variability. A rat does not always respond to food placed before it, and a factor called its "hunger" is invoked by way of explanation. The rat is said to eat only when it is hungry. It is because eating is not inevitable that we are led to hypothesize the internal state to which we may assign the variability. Where there is no variability, no state is needed. . . . In dealing with the kind of behavior that gives rise to the concept of hunger, we are concerned with the strength of a certain class of reflexes and with the two principal operations that affect it—feeding and fasting (22, pp. 341, 343).

For a considerable class of stimuli found to affect choice behavior in the maze, there is a fairly well demarcated class of events in the extramaze activities which exert an effect. Food, water, a female rat, all depend for their efficacy upon a deprivation schedule of some sort. For other stimuli, the rest of the day's activities seem of less relevance. For example, the effects of turning off a light in the Skinner box upon the lever pressing response would not depend upon a schedule of extra box illumination in any such obvious way as the effects of a food pellet depend upon the extra maze operations of feeding and fasting. Even here, at the extremes, it is likely that the schedule has some effect. Although I know of no experimental material on the point, it would be surprising if rats raised and maintained in a dark or extremely bright living cage would show the same response to light-off as a reinforcing agent. In order to keep the discussion quite general, I shall refer to *schedule-reinforcer* combinations, which will be understood to include those combinations in which almost any life-maintaining schedule is adequate. Whether there are any such does not need to be settled here. The stimulus presented is a *reinforcer*, and the presentation of it (an "event") is a *reinforcement*.

We are now in possession of a rather simple set of empirical facts. A certain stimulus, for a rat which has been under a specified schedule, for instance the sunflower seeds for a rat who has not ingested anything for 23 hours, will exert a strengthening effect. We can formulate a "law" stated crudely as follows: "In a rat which has not recently ingested sunflower seeds, bran mash, Purina chow, etc., a response of turning in a given direction in the T-maze will be increased if the fairly immediate presentation of sunflower seeds, etc., is made contingent upon that response." Similarly, we would find such a specific law to hold for thirst and water, sex and a mate, and so on. The general form of such special laws would be: "On schedule M, the termination of response sequence R, in setting S, by stimulus S^1 is followed by an increment in the strength of S.R." Such a law may be called a *situational-reinforcement* law, where the "reinforcement" is understood to stand for "presentation-of-a-reinforcer-following-a-specified-maintenance-schedule," and the term "situational" covers "response R in situation S."

Actually, in any given case, M, R, S, S^1 are classes. This is indicated by the suspicious-looking "etc." in the first "law" above. There is nothing shady about this "etc.," inasmuch as what is actually involved here is a class of operations and effects which are ultimately to be specified by locating each instance with respect to a whole complex set of dimensions. For example, Guttman (6) shows a relation between concentration of sugar solution used as a reinforcing agent and the strength of the lever pressing response. Heron and Peake (7) have studied protein as a specific component of reinforcement. There is to be discovered a vast number of such rather special laws which are comparable to the myriads of laws in chemistry concerning the solubility of substance Y in substance X and the like.

The next thing to notice is that while the schedule, reinforcement, response, and situation are all classes showing certain relations to one another, in general the schedule and reinforcer are related to one another more intimately than either is to the situation or response. The strength of a response which is maintained by food reinforcement is heavily dependent upon the feeding-fasting schedule, whereas the effect of a food reinforcement upon a response is relatively independent of, say recency of copulatory activity, so that a given schedule-reinforcement pair are "tied" to one another. But the Martian observes that the strengthening effect of a given schedule-reinforcement combination is relatively (not wholly!) neutral with respect to the response we are trying to strengthen and the situation in which we are trying to strengthen it. For a hungry rat, right turning depends heavily upon finding food; for a satiated rat, it depends very little. So the feeding schedule is intimately related to the reinforcing agent's efficacy. However, this "hungry-food" schedule-reinforcement combination seems to be capable of strengthening chain-pulling, lever-pressing, wheel-turning, marble-rolling, gnawing-through-paper, and so on through a very wide range of behaviors differing greatly in their topography and in their stimulus conditions. This leads to the question, will a certain schedule-reinforcer combination increase the strength of *any* response, in *any* setting?" This question turns out empirically to be answered in the negative, since we find at least three limitations upon the generality of a schedule-reinforcer combination as response strengthener. Leaving out the trivial case in which the response is anatomically impossible, e.g., to teach an elephant to thread a needle, we find:

1. No situation-response sequences may involve stimulus dimensions which are not discriminable by the organism. (Tolman's "discriminating capacities").
2. Some response sequences seem on the basis of their sequence, timing, or "complexity" not to be learnable by members of a given species, or subgroups within a species. It appears impossible to teach a rat a quintuple alternation problem, or to teach a human moron integral calculus.
3. There are cases in which the response we wish to strengthen is incompatible with responses at a very high (and relatively unmodifiable) strength under the schedule-stimulus combinations we are employing. For example, it would probably be next to impossible to teach a very hungry cat to carry a piece of fresh liver across the room, deposit it in a box, and return to receive food as a reinforcement. "Defensive" and "anxiety-related" responses are among the most important examples of this case.

How do we discover what responses have these characteristics? Experimentally, as we discover anything else. Let us call a situation-response combination having none of these properties *learnable*. A positive definition will be given below. What we find is that whereas learnable responses seem to differ somewhat in their "readiness" under different schedule-reinforcement combinations, this is a matter of parameters and does not invalidate the following tentative "law," which is stated qualitatively: "Any learnable response will be strengthened by sunflower seeds as a reinforcer." The general form of such a law is "the stimulus S^1 on schedule M will increase the strength of any learnable response." I shall call such a law a *trans-situational reinforcement* law. It must be noted carefully that such a law is still about a *particular* reinforcing agent, having, to be sure, a class character; but the particular reinforcing agent (and its associated necessary schedule, if any) is no longer tied to the response sequence first studied. The reinforcing property of sunflower seeds was noted first in the T-maze. The Martian will discover that white rats *can* learn to pull chains, press levers, and roll marbles. He finds that these learnable responses can also be strengthened by making the feeding of sunflower seeds contingent upon them. He makes the inductive generalization that sunflower seeds would exert this effect upon all learnable responses in the rat.

He now asks the obvious question: Are all schedule-reinforcer combinations like this? That is to say, when we study a new schedule-reinforcer combination and find it strengthens a response, can we assume that it will increase the strength of all learnable responses?

Naturally, our confidence in the general reinforcing power of any particular one will increase as we try it out on more and more learnable responses. But we do not know whether a higher-order inductive statement is justified, so long as we study sunflower seeds only or study several kinds of agents but in only one situation each.

Having found a particular reinforcer in a particular situation, we have discovered that it is trans-situational. Next we discover that all of the reinforcers that we have investigated have turned out to be trans-situational. The next induction is, "If a learnable response is followed by a stimulus which is known to be a reinforcer of learnable responses the strength will increase." A shorter way of saying this, having first defined a reinforcer as "a stimulus which will increase the strength of at least one learnable response," is simply: *all reinforcers are trans-situational*. Nothing is said as to the *amount* of strengthening. It is sufficient, in order to demonstrate the trans-situational character of a reinforcing agent, to show that it produces an increment in strength. If equal increments were required, it is probable that very few (if any) reinforcers would be trans-situational because of the varying behavior readinesses and different parameters of habit acquisitions from one drive to another and from one situation to another.

This assertion, that all reinforcers are trans-situational, I propose to call the *Weak Law of Effect*. It is not our problem in this paper to discuss whether the Weak Law of Effect holds or not. A "proof" of the Weak Law of Effect consists, as usual, of establishing inductively many instances of it in a variety of situations with our confidence increasing on the basis of the usual inductive canons. A "disproof" of the Weak Law of Effect would involve showing that a certain stimulus change acts as a reinforcing agent for one response, i.e., that the presentation of this stimulus following the response will increase the latter's strength; but that another response, previously established as learnable, cannot be strengthened by a presentation of this agent. A failure of the Weak Law of Effect to hold strictly would not be particularly serious, since one could (at the very least!) specify the exceptions and would hope to be able to generalize about them, that is, to discover empirically what are the kinds of reinforcers, or kinds of differences among situations, which reveal its invalidity. Actually, here again we have a case in which the law is stated in a qualitative all-or-none form; but the development of a science of behavior would eventually result in substituting a multiplicity of laws indicating the extent to which the reinforcing (strengthening) property generalized over various dimensions of the stimulus side, the reinforcing agent, and the "required" response properties. Assuming the Weak Law of Effect to have been established inductively, where are we now in our development? We have specific situation-reinforcer laws which state that a given stimulus is a reinforcing agent for a specified kind of response in a specified situation. As an example, we discover that for a standard rat, sunflower seeds will strengthen right turning in the T-maze. Having established several such specific situation-reinforcer laws, we find it convenient to introduce a definition, saying that a situational reinforcer is a stimulus which occurs as a term in such a specific situation-reinforcer law. Sunflower seeds are hence situational reinforcers. This definition is "arbitrary" or "conventional" in the usual sense, but clearly leads to no circularity. We cannot tell from the definition whether or not there is such a thing as a situational reinforcer, just as we cannot tell from the definition of a unicorn or of the phrase "King of France" whether such a thing exists. All we stipulate in the definition is that if a thing having certain properties turns out to exist, we will call it by this name. That there are situational reinforcers, that is to say, that we can find stimuli that do increase the strength of responses in a certain situation, is an empirical result. It is obvious that the specific situation-reinforcer laws have a perfectly good factual content (e.g., each such law could be false) in spite of the conventional character of the definition.

If our science contained nothing but a collection of such situational-reinforcer laws, we would still be in possession of valuable information. But we discover inductively that we can actually say more than this. For any given reinforcer, we discover that it can in fact be used to increase the strength of responses differing very greatly in topography from the one which originally led us to infer that it was a reinforcer, and in very different stimulating fields. It is true that there are a few special cases, as our cat with the liver, in which we cannot increase the strength of a *kind* of a response (carrying an object from one place to another) which we know from independent study this species is able to learn. But in all such cases we are able to specify an interfering response at such high strength that the behavior in question does not get a chance to be emitted, and hence cannot be reinforced. With this exception, we are able to say that a given reinforcer will increase the strength of all learnable responses of the species; although there will be quantitative differences which remain to be discovered and generalized about after much painstaking experimentation. We define a reinforcer which is of this sort as trans-situational, and from a study of numerous reinforcers we conclude that they are all of this type. The second order induction that all reinforcers are trans-situational (the Weak Law of Effect) is then made.

This last is certainly a very rich and powerful induction. It is true that to make predictions we must study at least one learnable response in order to find out whether a given stimulus change is reinforcing, and we must know for any contemplated response whether it is learnable. Experience with a given species need not be too extensive in order to get a general idea of the kinds of behavior which are possible and learnable; and once having this, we proceed to strengthen responses by means of reinforcing agents which have never been utilized before in connection with these responses. This is so commonplace that we are likely to underestimate its theoretical significance. So far as I know, no animal psychologist has the least hesitation in utilizing any of a very large class of reinforcing objects called "food" in experimentation upon practically any kind of behavior which he is interested in studying. Should he find a failure of response strength to increase, the chances of his asking what is wrong with the food are negligible. His inductive confidence in the Weak Law of Effect is such that he will immediately begin to investigate what is wrong with the stimulus field, or what requirements concerning the response properties he has imposed which transcend the powers of the organism. I am stressing this point because there is a tendency to say that since we have to study the effects upon strength in order to know whether an agent is reinforcing, we do not really "know anything" when we have enunciated the Law of Effect. I think it should be obvious from the diversity of both the class called learnable and the class of agents called reinforcing that to the extent that this law holds almost without exception, when we have enunciated it we have said a great deal.

The man from Mars might be tempted here to take a final step which would be suggested by the ubiquity of the manifestations of the Weak Law of Effect. It might occur to him that the great majority of the instances in which changes in response strength occur seem to involve the operation of the Weak Law, i.e., the presentation of a member of the reinforcing class. Perhaps it is not only true that any learnable response can be strengthened by the presentation of a trans-situational reinforcer but may it not be that this is the *only* way to increase the strength of responses (by learning)? Response strength may be increased by surgical and drug procedures, and also by maturation; but the demarcation of learning as a very general mode of response change, while it presents difficult problems, need not concern us here. Assuming that we have some satisfactory basis for distinguishing an increase in the strength which is based upon "experience" rather than upon interference with the reaction mechanism or biological growth determined by genetic factors given minimal (viable) environments, we may ask whether learning takes place on any *other* basis than the Weak Law of Effect. Certain apparent exceptions to this statement of reinforcement as a necessary condition would appear, but the Martian might ask whether these exceptions are more apparent than real. The formulation of such a law would run something like this: "Every learned increment in response strength requires the operation of a trans-situational reinforcer." I shall designate this rash inductive leap as the *Strong Law of Effect.*

It appears obvious that this also is a statement far from being experimentally empty or in any sense a consequence of definition. I have heard psychologists translate the statement "he learns because he was reinforced" as being tantamount to "he learns because he learns." Postman suggests the same kind of thing in the first quotation above. This is too easy. The expanded form which I suspect everyone has implicitly in mind when he talks about the Strong Law of Effect is: "He learns following the presentation of a stimulus change which for this species has the property of increasing response strength; and, other things being equal in the present setting, if this change had *not* occurred he would not have learned." Such a statement can clearly be false to fact, either because no such trans-situational reinforcer can be shown to have been present, or because the same learning can be shown to be producible without it in the present setting. The claim of the reinforcement theorist to explanation is (at this stage of our knowledge) of exactly the same character as "he developed these symptoms because he was invaded by the Koch bacillus, and we know that the Koch bacillus has these effects." This is not a very *detailed* explanation, because the intermediate or micro-details of the causal sequence are not given; but it is certainly neither factually empty nor trivial.

In our initial quotation from Postman, we find him saying, "The satisfying or annoying nature of a state of affairs can usually be determined fully only in the course of a learning experiment and cannot then be invoked as a causal condition of learning without circularity." The trouble with this remark lies in the ambiguity of the phrase "*a* learning experiment." That we cannot know what is reinforcing without having done *some* experimentation is obvious, and is just as it should be in an empirical science. But once having found that a certain state of affairs *is* reinforcing for a given species, there is no reason why a given case of learning cannot be explained by invoking the occurrence of this state of affairs as a causal condition. The definition of force does not entail the truth of Hooke's law. It is only by an experiment that we find out that strain is proportional to stress. Once having found it out, we are all quite comfortable in utilizing Hooke's law to account for the particular cases we come across. I am confident that Postman would not be disturbed if in answer to the question, "Why does that door close all the time?" someone were to reply, "Because it has a spring attached to it on the other side." There is no more "circularity" in this kind of causal accounting than in any other kind. It is perfectly true that this kind of "lowest-order" explanation is not very intellectually satisfying in some cases, although even here there is a considerable variability depending upon our familiarity with the situation. For a detailed consideration of these problems by more qualified persons I refer the reader to papers by Hospers (**10**), Feigl (**4, 5**), and Pratt (**20**).

I think it is obvious that this is the way we think of the Law of Effect, whatever we may think as to its truth. When an apparent case of learning in the absence of reinforcement occurs, those who are interested in preserving the status of the Law of Effect (in my terminology, in preserving the status of the *Strong* Law of Effect) begin to search for changes following the response which can be shown to be of the reinforcing sort. They do not simply look for *any* stimulus change and insist ("by definition") that it is a reinforcement. The statement that a given case of apparently non-reinforcement learning is actually based upon secondary reinforcement is essentially a claim that some stimulus change can be shown to have followed the strengthened response, and that this stimulus change has (still earlier) been put in temporal contiguity with a stimulus change of which we know, from a *diversity* of situations, that it exerts a reinforcing effect.

Abandoning the charge of circularity, a critic might offer a "practical" criticism, saying, "What good does it do to know that a reinforcer strengthens, when the only way to tell when something is a reinforcer is to see if it strengthens?" The trouble here lies in the vagueness, since the *generality* is not indicated, and this failure to indicate generality neglects the usual advantages of induction. That a describable state of affairs *is* reinforcing can only be found out, to be sure, by experimenting on some organisms utilizing *some* learnable response. But it is not required (if the Weak Law of Effect is true) that we, so to speak, start afresh with each new organism of the species and each new response. As a matter of fact, after we have considerable experience with a given species, we can generalize about the physical properties of a stimulus class. So that finally "food" means many substances which may never yet have been tried in a learning situation, and may never have been presented in natural circumstances to the members of a particular species. Wild rats do not eat Purina Chow. Here we begin to approach inductively one of the previously rejected bases of defining reinforcement, namely, the physical character of the stimulus change itself. To ask for a definition of reinforcers which will tell us beforehand for a given species which objects or stimuli will exert the reinforcing effect is to ask that a definition should tell us what the world is like before we investigate it, which is not possible in any science. It happens that the psychologist is worse off than others, because species differences, individual hereditary differences, and differences of the reactional biography make a larger mass of facts necessary in order to know whether a given agent will reinforce a particular organism. But at worst the Weak Law of Effect in conjunction with its member laws is far from useless. When I know inductively that all non-toxic substances containing sugar will act as reinforcers for organisms from rat to man and therefore that I can almost certainly strengthen all responses learnable by any of these species on the basis of the presentation of any of these substances, I know a great deal and my science has a very considerable predictive power.

AN ANALOGOUS PROBLEM IN PHYSICS

It is instructive to consider a somewhat analogous problem in physics, in the definition of "force." Once mass has been defined by some such artifice as Mach's acceleration-ratio technique, and acceleration defined in terms of time and distance, Newton's second law is a *definition* of force. I neglect here other attempts to introduce the notion such as the "school of the thread" (**18**), utilizing Hooke's law in the form of a definition rather than a law, or its modern variants, e.g., Keenan's (**13**) recent effort. Force is "that which accelerates mass." Mach's introduction of the concept of mass was somewhat disturbing to certain of his contemporaries because of a suggested circularity. Mach saw that it was the *inertial* character of mass, rather than "weight" or "quantity of matter" which was crucial in setting up the definition of force. Accordingly, he proceeds as follows:

> a. *Experimental Proposition.* Bodies set opposite each other induce in each other, under certain circumstances to be specified by experimental physics, contrary *accelerations* in the direction of their line of junction. (The principle of inertia is included in this.)
>
> b. *Definition.* The mass-ratio of any two bodies is the negative inverse ratio of the mutually induced accelerations of those bodies.
>
> c. *Experimental Proposition.* The mass-ratios of bodies are independent of the character of the physical states (of the bodies) that condition the mutual accelerations produced, be those states electrical, magnetic, or what not; and they remain, moreover, the same, whether they are mediately or immediately arrived at.
>
> d. *Experimental Proposition.* The accelerations which any number of bodies A, B, C. . . . induce in a body K, are independent of each other. (The principle of the parallelogram of forces follows immediately from this.)
>
> e. *Definition.* Moving force is the product of the mass-value of a body into the acceleration induced in that body. Then the remaining arbitrary definitions of the algebraical expressions "momentum," "vis viva," and the like, might follow. But these are by no means indispensable. The propositions above set forth satisfy the requirements of simplicity and parsimony which on economico-scientific grounds, must be exacted of them. They are, moreover, obvious and clear; for no doubt can exist with respect to any one of them either concerning its meaning or its source; and we always know whether it asserts an experience or an arbitrary convention (**17**, pp. 243–244).

In the appendix to the second English edition, Mach replies to critics of this procedure as follows:

> A special difficulty seems to be still found in accepting my definition of mass. Streintz has remarked in criticism of it that it is based solely upon gravity, although this was expressly excluded in my first formulation of the definition (1868). Nevertheless, this criticism is again and again put forward, and quite recently even by Volkmann. My definition simply takes note of the fact that bodies in mutual relationship, whether it be that of action at a distance, so called, or whether rigid or elastic connexions be considered, determine in one another changes of velocity (accelerations). More than this, one does not need to know in order to be able to form a definition with perfect assurance and without the fear of building on sand. It is not correct as Höfler asserts, that this definition tacitly assumes *one and the same force* acting on both masses. It does not assume even the notion of force, since the latter is built up subsequently upon the notion of mass, and gives then the principle of action and reaction quite independently and without falling into Newton's logical error. In this arrangement one concept is not misplaced and made to rest on another which threatens to give way under it (**17**, pp. 558–559).

It is obvious that Mach defines mass in the way he does *so that* the definition of force by $F=ma$ will lead to the kinds of laws we want. That is, a previous "knowledge" of the law of gravity based upon a cruder notion of mass is involved historically in the formulation of such a definition. But the crucial point is that it is involved only in the context of discovery, not in the context of justification (**21**, pp. 6–7). There is nothing wrong with making use of any notions, including vague anthropomorphic experiences of pleasure, in deciding how we shall formulate definitions, since our aim is to erect concepts and constructs which will fit into the most convenient and powerful system of laws. The point is that we wish to come out with explicit notions that are free of this vagueness and which do not require any notions which cannot be finally introduced objectively. There is probably a remnant of hedonism in the thinking of the most sophisticated contemporary reinforcement theorists, and there is no reason why anybody should pretend that when he talks about rewards he does not have some faint component in his thinking which involves the projection of such pleasure-pain experiences. But this does not mean that these notions are made part of the scientific structure he erects, in the sense that either the definitions of terms or the establishment of the laws requires such associated imagery in his readers. I suggest that Thorndike's critics are in the same position as Mach's.

One might ask, why would a physicist be upset should he attend a spiritualist seance and find tumblers leaping off tables and floating through the air? If the concept of force is given simply by the relation

$F=ma$, then, if a glass tumbler undergoes an acceleration, a force must act and his definition assures him that the physical world will not surprise him. I do not think the answer to this question is far to seek. While it is admittedly a question of decision, I doubt that most physicists would decide to say that an acceleration occurred in the absence of a force. If the genuineness of the phenomenon were satisfactorily established, I do not think there would be a re-definition of the *concept* of force, but rather that the existence of "forces" on other bases than those previously known would be assumed. That is, the physicist would not say "here is a case of acceleration without a force," but he would rather say "here is a case of force not arising from the usual mechanical, gravitational, or electro-magnetic situations which I have thought, up to now, were the sole bases on which forces came into being." It is certainly no criticism of a Newtonian definition of force (I leave out the fact that Newton, while he defined force in this way, apparently also treated his second law as one of empirical content) to say that having thus defined force you cannot know beforehand what are the conditions in the world under which forces will appear. The mechanical forces involved in direct contact, the force of gravity, and certain electrostatic and magnetic forces were known to Newton. There is nothing about his definition of force which tells us that a peculiarly directed force will exist between a wire carrying an electric current and a compass needle, nor that attracting or repelling forces will exist between parallel wires each of which carries a current. The discovery of these conditions under which forces exist was an empirical contribution of Oersted and Ampere.

Similarly, the psychologist defines what is meant by a reinforcer, and proceeds to search for the agents that fall under this definition. There are undoubtedly kinds of stimulus changes of which we are as yet unaware which will turn out to have the reinforcing property. Dr. Wilse Webb (personal communication) has found in preliminary experiments that at least in one kind of Skinner box the click produced by the operation of an empty magazine will exert a reinforcing effect in an animal whose experience has never given this stimulus an opportunity to acquire secondary reinforcing properties. This is surprising to us. What are the conditions under which this will occur? Suppose it should be found that almost *any* stimulus change within a fairly wide range (avoiding extreme intensities which are anxiety-producing) would exert a slight reinforcing effect in the Skinner box or in any similar apparatus in which there is a considerable stimulus restriction and a marked constancy in the homogeneity of the visual and auditory fields. It might be discovered that when a member of this species has remained in such a homogeneous field for a period of time, stimulus *changes* (not otherwise specified) exert a reinforcing effect. Maybe the rat is "bored" and just likes to make something happen! A difficult notion to nail down experimentally, to be sure. But its complexity and the number of things to be ruled out, does not take it out of the realm of the confirmable.

Let us consider a very extreme case. Suppose in the T-maze situation a systematic increase in the strength of the right turn should be discovered for a standard rat. Suppose that the most thoroughgoing, exhaustive manipulation of the external effect of right-turning should fail to reveal any condition necessary for the effect. "No member of the reinforcing class is to be found." I think that at this point we would begin reluctantly to consider a reinforcing property of the response itself. Perhaps turning to the right is inherently reinforcing to this species. It seems, for instance, that "fetching" behavior in certain species of dogs is self-reinforcing (or at least that it has a biologically replenished reserve). The only reason for calling right-turning "self-reinforcing" rather than simply saying that it is a response of innately high strength in the species is that a *change* in strength occurs with successive runs, otherwise "turning to the right" is simply a kind of tropism. Is the "self-reinforcing" idea factually empty? Although many people would disagree with me at this point, I do not think it is. But it has factual meaning only intradermally. There is no reason why we could not study the proprioceptive effects of a right turn and find out whether, if they are cut out, the increase in response strength continues to occur. In principle we could create the proprioceptive effects of a right turn by artificial means and on that basis strengthen a topographically different response such as lifting the fore paw, wiggling the whiskers, or the like. Here there are difficulties, but I would be prepared to argue that in principle the self-reinforcing effect of right-turning is an empirically meaningful notion.

An interesting side-light is that even the Strong Law of Effect is, as stated, compatible with the latent learning experiments. I am not interested in avoiding the consequences of those experiments by shrewd dialectics, but in the interests of clarity it should be pointed out that in, e.g., the Blodgett design, the big drop in errors *does* follow a reinforcement. So long as the Strong Law of Effect is stated qualitatively and does not explicitly mention amounts and times, it would be admittedly difficult to design an experiment in which it could be refuted. A neo-Hullian interested for some reason in preserving the Strong Law of Effect might simply add a quantitative postulate. He might assume that when a response undergoes an increment in strength on the basis of a minimally reinforcing agent (that is, one in which the asymptote of the acquisition of habit strength is relatively low), then, if subsequently a strong reinforcement is introduced, the parameter in the new growth function which determines the rate of approach to the new asymptote is greater than it would have been without the original learning. Since in the Blodgett design there is evidence of acquisition of differential habit strengths during the latent phase, such a postulate would lead to a preservation of the Strong Law of Effect. The main reason that we are concerned to deal with latent learning material of the Blodgett type is that in the reinforcement theory as now formulated, the effect of a reinforcer is implicitly assumed to operate immediately.

RELATIONSHIP OF REINFORCEMENT TO DRIVE

Perhaps a comment is needed on the way in which reinforcement has been treated here as the primary notion whereas drive, need, or demand is defined in terms of it. I do not mean to imply that need or drive is not the more "basic" factor, if by this is meant that what is a reinforcer or what acquires reinforcing properties depends upon a certain relevance to need. But this manner of speaking refers to the causal reconstruction of behavior, and reverses the epistemological order. The needs of an organism are inferred from changes in behavior strength as a function of certain states of affairs. That is to say, we "get a fix" on a need by being able to induce the chief defining properties of those states of affairs to which behavior is shown to tend. I do not see how there is any possibility in proceeding otherwise at the level of molar behavior. Whether it will be feasible or desirable to hypothesize a kind of state called need in the case of all reinforcers is a moot point at present. I gather that Hull would argue it will, whereas Skinner would argue it will not. One can consider a sort of continuum of reinforcing states of affairs at one end of which it is most easy and natural and obviously very useful to speak in terms of a need, e.g., the case of food or water; whereas at the other end, e.g., the reinforcing effects of hearing a click or turning off a light, the notion of needs seems relatively less appropriate. But the *causal* primacy of needs in our final reconstruction of behavior laws must not be confused with the epistemological status of needs, i.e., the operations by which we arrive at a conception of the needs. Whether the reduction of need is a necessary condition for learning is a question that is not involved in my formulation of either the Weak or the Strong Law of Effect since need-reduction is not equated to reinforcement. This independence of the notions of rein-

It seems clear that in the above sequence both the definitional and the factual (empirical) elements are present, and in a simple, commonplace form. The definitional and conventional elements appear in the specification of the circumstances under which a stimulus is to be called "reinforcing." Such a stipulation, however, cannot tell us whether any such stimuli exist. That they do exist, which no one doubts, is an empirical finding; and the numerous statements about them constitute situational-reinforcer laws which are in a sense the special "sub-laws" of effect. These are related to the Weak Law of Effect somewhat in the same way that the particular empirical laws about the properties of bromine, fluorine, chlorine, and so on, are related to the Periodic Law. That the stimuli which occur in the situational-reinforcer laws have a generality of their reinforcing power is also an empirical finding, at present less well established (the Weak Law of Effect). That all cases of learning require certain time relationships to the presentation of such general reinforcers is yet a further factual claim, at present very much in dispute (the Strong Law of Effect).

I can see no reason why any theorist, whatever his position, should find the preceding treatment objectionable as an explication of the Law of Effect. I do not see any way in which the Strong Law of Effect, which is after all the big contemporary issue, has been surreptitiously put into the definitions in such a way that what is intended as an empirical proposition is effectively made a consequence of our use of words. The status of the Strong Law of Effect and even to some extent the Weak Law is presently in doubt. Further, some of the words used in these definitions, e.g., the word "response," are difficult to define in a way that makes them behave in the total system as we wish them to. I have not tried to deal with all these problems at once, but I hope that there are no difficulties springing from the problem of circularity which have not been met. That it is difficult to untangle the learning sequence which has given the reinforcing property to some states of affairs, particularly in the human organism, is admitted by everyone. That a large amount of detailed work of the "botanizing" type, not particularly ego-rewarding, needs to be done before the special sub-laws of effect are stated in terms of quantitative relations is quite clear. Finally, it would be very nice if in some magical way we could *know* before studying a given species exactly what stimulus changes would have the reinforcing property; but I have tried to indicate that this is an essentially irrational demand. In the light of the previous analysis I think the burden of proof is upon those who look upon a sophisticated formulation of the Law of Effect as circular, in either of the ordinary uses of that word.

BIBLIOGRAPHY

1. CARR, H. A., *et al.* The Law of Effect: a roundtable discussion. *Psychol. Rev.*, 1938, **45**, 191–218.
2. ESTES, W. K. A study of motivating conditions necessary for secondary reinforcement. *Amer. Psychologist*, 1948, **3**, 240–241. (Abstract.)
3. FARBER, I. E. Response fixation under anxiety and non-anxiety conditions. *J. exp. Psychol.*, 1948, **38**, 111–131.
4. FEIGL, H. Operationism and scientific method. *Psychol. Rev.*, 1945, **52**, 250–259.
5. FEIGL, H. Some remarks on the meaning of scientific explanation. In H. Feigl & W. Sellars., *Readings in philosophical analysis*. New York: Appleton-Century-Crofts, 1949. Pp. 510–514.
6. GUTTMAN, N. On the relationship between resistance to extinction of a bar-pressing response and concentration of reinforcing agent. Paper presented at the meeting of the Midwestern Psychological Association, Chicago, Ill., April 29, 1949.
7. HERON, W. T., & PEAKE, E. Qualitative food deficiency as a drive in a discrimination problem. *J. comp. physiol. Psychol.*, 1949, **42**, 143–147.
8. HILGARD, E. R. *Theories of learning*. New York: Appleton-Century-Crofts, 1948.
9. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning*. New York: Appleton-Century, 1940.
10. HOSPERS, J. On explanation. *J. Philos.*, 1946, **43**, 337–356.
11. HULL, C. L. Thorndike's *Fundamentals of learning*. *Psychol. Bull.*, 1935, **32**, 807–823.
12. HULL, C. L. *Principles of behavior*. New York: Appleton-Century, 1943.
13. KEENAN, J. Definitions and principles of dynamics. *Sci. Mon., N. Y.*, 1948, **67**, 406–414.
14. LENZEN, V. F. *The nature of physical theory*. New York: John Wiley, 1931.
15. LINDSAY, R. B., & MARGENAU, H. *Foundations of physics*. New York: John Wiley, 1936.
16. MACCORQUODALE, K., & MEEHL, P. E. "Cognitive" learning in the absence of competition of incentives. *J. comp. physiol. Psychol.*, 1949, **42**, 383–390.
17. MACH, E. *The science of mechanics* (Transl. by T. J. McCormack). Second English Ed. Chicago: Open Court Publishing Co., 1902.
18. POINCARÉ, H. *The foundations of science*. New York: Science Press, 1913.
19. POSTMAN, L. The history and present status of the Law of Effect. *Psychol. Bull.*, 1947, **44**, 489–563.
20. PRATT, C. C. Operationism in psychology. *Psychol. Rev.*, 1945, **52**, 262–269.
21. REICHENBACH, H. *Experience and prediction*. Chicago: Univ. of Chicago Press, 1938.
22. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
23. SPENCE, K. W. Studies on secondary reinforcement. Address given to the Minnesota Chapter of Psi Chi, Minneapolis, April 22, 1948.
24. TAYLOR, L. W. *Physics, the pioneer science*. New York: Houghton, Mifflin, 1941.
25. THORNDIKE, E. L. *The fundamentals of learning*. New York: Teachers College, Columbia Univ., 1932.
26. THORNDIKE, E. L. *Animal intelligence*. New York: Macmillan, 1911.
27. THORNDIKE, E. L. *The original nature of man*. New York: Teachers College. 1913.
28. THORNDIKE, E. L. *The psychology of learning*. New York: Teachers College, 1913.
29. TOLMAN, E. C. *Purposive behavior in animals and men*. New York: Appleton-Century, 1932.

Received July 27, 1949.

THORNDIKE, ROBERT L. *Personnel selection test and measurement techniques.* New York: John Wiley, 1949. Pp. viii+358. $4.00.

In one sense, the health of a science may be assessed upon the basis of its ability to advance when granted unusual opportunity, and the manner in which that growth is reflected in its texts and reference volumes. Psychologists rose to the challenge and opportunity given them in the field of selection by the war and demonstrated that they could perform a useful function while increasing their fund of knowledge and understanding. Now Thorndike has shown that the results of this work can be brought to the student or layman in a well-organized, concise, and internally consistent account of basic principles and techniques.

Indeed, this book fills a need which has been rapidly approaching the acute. The confusion and downright misinformation which mark the discussion of selection testing in most of our textbooks have done much to encourage the tendency toward test validation by analogy or fiat. This tendency is now appearing in the work of industrial, clinical, and guidance psychologists alike. The author's powerful presentation of basic theory, accompanied by his consistent acceptance of the demands of logic and scientific method, should do much to clear the air of the fallacious and the opportunistic.

The first section of the book is devoted to the techniques of building and evaluating a test selection program. Beginning with a chapter in which job analysis is properly recognized as ". . . . a source of hypotheses for selection tests and a source of insight about criterion measures" (p. 13), Thorndike proceeds to a discussion of various types of selection instruments and their construction. This discussion is sound and comprehensive. It is in the following chapters on reliability and validity, however, that the author makes new and important contributions. The discussion of reliability is pursued with originality and a firm refusal to oversimplify what is really a complex and vitally important concept. The conclusion that "It becomes clear that the basic problem for determining the reliability of any measurement procedure becomes that of *defining* what shall be considered true variance between individuals and what shall be thought of as error variance" (p. 72), and the development which follows should serve to make many who believed they had mastered the concept approach it with a new respect.

The chapters on validity are of equal stature. The problem of criteria, shamefully neglected for so long by so many, at last receives a precise and careful analysis. Not the least of the author's virtues is his ability to unearth and illuminate unsolved or difficult problems. When the reader has finished following Thorndike's consideration of criteria, he cannot fail to recognize at one time how stringent are the requirements for satisfactory performance measures and with what difficulty they are met. On the subject of the estimation of test validity, there is, once again, a skillful dissection of many of the complexities which have been skirted in most texts. Having made a point which is only too often ignored in the interests of convenience, "Validity coefficients based on fewer than 100 cases are too unstable to serve as the basis for any serious comparative study of different tests for personnel selection . . . ." (pp. 58–59), the author considers validity estimation under various combinations of criterion and test score with regard to the factor of continuity or discreteness. The assumptions underlying the use of various correlational measures are clearly stated and the pitfalls encountered in their interpretation receive the appropriate danger signals. Of particular interest is the treatment of population curtailment—an ubiquitous but neglected difficulty.

Chapters on the combination of tests into a battery and the analysis of test items include an evaluation of the clinical method of test combination and interpretation. The author's restrained but firm insistence upon the fact that the clinical method is, of its nature, specific to the clinician and the personnel situation in comparison to the generality of the objective method of linear combination stands in clear contrast to the fuzzy thinking which has resulted in attempts to distinguish between empirical validity and "clinical validity," whatever the latter term may mean.

The author then proceeds to a very practical consideration of the administration of testing programs. This account, covering such topics as the conduct of testing, scoring and weighting procedures, the organization of reports and records, quotas and yields, and the effective public presentation of test research and results, should serve two important functions. For the student, these chapters should help to bring recognition that personnel testing is not performed in a vacuum but is, instead, applied to real people doing everyday jobs. For the personnel executive and for the psychologist whom he employs, the chapters should produce a respect for the necessity to understand what functions the testing program is to perform and how it may be designed to fulfill them best. In this connection, the author's discussion of the relationship between selection, multiple selection, and classification is particularly useful.

Of course, the book has some faults. It is probable that the author would be the first to admit that he has not solved the ever-present difficulty of insuring that the reader will understand the statistical concepts so necessary to a grasp of this material. His implication that a background corresponding to about a year of work in statistics is sufficient for his readers seems unrealistic to this reviewer. Students will need help and plenty of it in most of the chapters. The writing is clear and readable, but there has been no attempt to predigest or to prettify.

The reader is expected to work and so, quite properly, is the instructor.

The wide adoption of this work as a text in courses on personnel seems inevitable. If those who test and interpret also adopt the integrity of the book's approach and its basic philosophy, psychology will be considerably in Thorndike's debt.

S. RAINS WALLACE, JR.

Life Insurance Agency Management Assn.

CHASE, STUART. *The proper study of mankind. An inquiry into the science of human relations.* New York: Harper, 1948. Pp. xx+311.

The books of Stuart Chase are neither popularizations nor original contributions to knowledge and science. Chase is a journalist of highest character and integrity in the fields of social science. In earlier books he has reported on subjects ranging from the waste of natural resources to semantics. The purpose of his latest book, he tells us, is "to run a kind of chain and compass line across the whole front of the sciences devoted to human relations." This is a formidable undertaking. The result does not wholly satisfy Mr. Chase nor is it likely to satisfy the specialist in any of the fields involved. But we may be grateful for the attempt.

The book is organized into three parts. Part I develops the conception of science as it is generally accepted today and illustrates the contention that social relations are subject to investigation by the scientific method. The point of view has been well developed before, but this is as clear and simple a statement of it as this reviewer has read. It is particularly effective in making the point that social science already has a small but significant body of empirically established generalizations "in the storehouse" which have been useful, and can be further useful, in application to problems of social relations. Chase debunks the idea that scientists are magicians suddenly producing an atomic bomb or a "grand solution." But the air of breathless suspense with which some of the researches are reported presents an equally unrealistic description of the painstaking little steps by which a science is built.

Part II, entitled Landmarks and Achievements, presents selected researches and key ideas primarily developed out of the cross-fertilizations of sociology, social anthropology, and social psychology. It also presents some comments on the dismal shortcomings of economics and political science as sciences. There are chapters summarizing Leighton's report on Poston, the current ideas on race and inter-group relations, Ogburn's theory of cultural lag, the social anthropologists' work on American communities, the Western Electric studies, and public opinion polls. These, and some others, are discussed with skill but with the inevitable over-simplification required to make them "landmarks." The integration of the section suffers from the fact that such researches as Kinsey's on the sexual behavior of American males and the strike in the shoe factories of Yankee City are not parts of an integrated theory of social behavior. This is not Mr. Chase's fault.

In Part III Chase attempts to pull the results of his summary together and point up the moral of his inquiry. Here he discusses communication as a basic problem with a separate chapter on the contributions of semantics. He cautions reformers to reform if they are to lay the scientific foundations for achieving one world. He points the "roads" to answers to practical problems: a gigantic project to plan for permanent peace, more teamwork, encouragement of social science through better recruitment of talented young people, more funds, and more public awareness.

The practical problems offered in this section as ultimate objectives of social science are usually stated in economic or political terms. But economics and political science are the very fields which least fulfill Mr. Chase's conception of science. This is partly because these fields lack the underpinning of the other social sciences and partly because they fail to use what is available to them. The basic reason is, however, that most economists and political scientists ask questions not of science but of policy.

Political science and economics are the senior members of the social science family but they have been the last to turn away from the problems of what *should* be. Perhaps their older emphasis should continue. Understanding of behavior—including economic and political behavior—can be approached through the sciences of human relations, but the directions of public policy must be derived from other sources. There is a place—possibly within a broadened conception of social science—for the disinterested consideration of the ends of social policy to which the instruments and findings of science can be put. Chase might well have considered this orientation as an alternative one for economics and political science.

Mr. Chase is not equally critical of all the social sciences. The shining examples of research useful in practical situations seem partially to blind him to the confusions of terminology and the shortcomings of systematic theory in social psychology, sociology and anthropology. Perhaps the general public to whom the book is primarily addressed needs such over-statements tempered only by the qualifications which Mr. Chase includes. Specialists in these fields will be inclined to think that their sciences have almost been over-sold. Despite the reminders from time to time that social science is in the early stages of its career, the net impression which the book is likely to give to the uninitiated reader is that social science is ready for application to many problems today. Actually it can make only limited contributions. It is no service to social science to build up expectations for its performance which it can only partially fulfill.

The book has been offered by the publisher as a textbook for college courses. In this role it falls short because of the very characteristics which make it a good book for the trade. It is neither systematic nor detailed enough to serve as an adequate summary of the sciences of human relations. It will, however, make excellent collateral reading. No other book has come to the reviewer's attention which gathers together in one place such lucid reports of good current research in all the major social sciences. For this reason, if for no other, it will serve a useful function for the departmentalized college student.

HENRY J. MEYER.

New York University.

SMITH, GUDMUND. *Psychological studies in twin differences.* Lund, Sweden: Gleerup, 1949. Pp. 251.

This is essentially a monograph reporting the author's investigations on 61 pairs of like-sex twins ranging in age from 8 to 69 years and living in the vicinity of Lund, Sweden. The twins were classified as identical or fraternal on the basis of birth records as well as a series of anthropometric characteristics including fingerprints, blood groups, iris color and structure, hair and skin coloring, and facial and cranial measurements. The total group contained 32 identical pairs (13 male and 19 female) and 29 fraternal pairs (16 male and 13 female); these numbers include one set of triplets who were entered as one identical and two fraternal pairs in most of the calculations.

Two parallel approaches were followed. One included the administration of imagery tests, including measures of duration, periodicity, and size of a negative after-image when projected at different distances, and an intensive study of eidetic imagery. Evidence of the latter was found in 25% of the subjects under 16 years of age and 9% of the older subjects. The second approach comprised a personality study, including two questionnaires answered by the twins themselves and two by their families, supplemented in some of the cases by teachers' or employers' ratings, school records, and a "pair test" in which a difficult construction puzzle was attempted first by each twin separately and then by the two working together, their reactions under both conditions being observed by an assistant. The personality data and the more qualitative results of the imagery tests are reported principally in the form of case studies of each pair, while the quantitative data are treated in terms of analysis of variance and intra-class correlation. It should be noted that the number of twin pairs on whom these various techniques were employed varied; the imagery tests, which constituted the main part of the study, were given to only 22 identical and 26 fraternal pairs.

Throughout the study, special attention was given to intra-pair relations and the specialization of roles. The relations of identical twins seemed to be characterized more often by coordination and cooperation, those of fraternal twins by competition and conflict. The influence of such factors as age, sex, intellectual level, and personality characteristics upon twin relationships is also discussed in detail. A number of imagery characteristics yielded significantly higher intra-pair correlations among identical than among fraternal twins. The investigator points out repeatedly that the higher concordance of identicals was more clearly evident in certain relative indices than in isolated measures of reaction. For example, the absolute size of the images showed no greater correlation among identical than among fraternal twins, but change in image size with repetition did show closer resemblance among identicals. Another example of these relative measures, which the author terms indices of "development" of the image, is based on the effects of "intermittence" upon size, duration, and periodicity of the image. Such "intermittence" was achieved by inserting a rotating disc with two black sectors in front of the eye-piece, while the subject was projecting the image. The investigator also purports to show that these measures of "development," as contrasted to isolated static scores, are more likely to correlate with basic personality characteristics. For a demonstration of this relationship, the author turns to the case study material on the same subjects.

On the basis of the above findings regarding the possible value of measuring *changes* of response under varying conditions, the author concludes that "every test should be a longitudinal study, however brief" (p. 233). This emphasis is probably the most suggestive contribution of the monograph to general experimental design. With respect to specific findings, two cautions need to be observed in interpreting the data. First, the number of cases on which any one comparison was based was often very small. Secondly, a strong element of subjectivity entered into most of the observations, especially those dealing with broad personality characteristics. Thus in the "pair test," no quantitative records are given, but the observations of the assistant have simply been incorporated into the general case reports. The questionnaires were used primarily as a basis for informal questioning by the experimenter, and again no specific replies are reported. Even more subjective is the subsequent analysis and classification of case reports and the linking of personality patterns with imagery characteristics. The role of experimenter's set cannot be discounted under these conditions. Finally, it might be added that the style and organization of the monograph make it very difficult to read. The prevalence of literal translation produces some extremely awkward passages as well as a very unusual use of English words. For example, the word "formula" as herein used signifies "questionnaire" and case studies are described as "casuistics." The entire presentation could have been considerably shortened and clarified.

ANNE ANASTASI.

Graduate School, Fordham University.

JOHNSON, DONALD M. *Essentials of psychology.* New York: McGraw-Hill, 1948. Pp. xiii+491. $3.50.

For many years, text-books designed for use in the beginning course in psychology have shown a tendency toward expansion. The author of the present book has reversed this process and produced one of the shortest texts for the beginning course that has been published in recent years. Of the 491 pages in the book, 442 contain the actual text, the remaining pages being devoted to references, a glossary, a list of audio-visual aids, and index. That the general content of the beginning course is covered in this short space is indicated by the chapter contents: motives and emotions, sensory functions, attention, perception, learning, thought and judgment. A little over half of the book (228 pages) is devoted to the preceding topics. The remaining 214 pages are divided into four chapters on the individual and society, abilities and tests, personality, and abnormal personalities. There is a summary at the end of each chapter and a list of technical terms for study. The book contains a total of 122 figures relevant to the text discussion. In general, the style is readable, concise, and informal.

The guiding principles which have aided the author in producing a text as short as this are described in the Preface. The importance of integration is stressed and principles which are referred to later in the text are therefore placed in the earlier chapters. Historical material and discussions of psychological schools are omitted. No attempt is made to document and prove every point. Classifications (of methods, behavior, etc.) are omitted unless they are used at least twice in later exposition. The author's hope that a book guided by these rules would be shorter than usual has been realized.

With a text as short as the present one, the obvious question arises: Is it possible to present the essentials of psychology adequately in 442 pages? The answer to this question will undoubtedly vary with the individual who reads the book. It is possible to find sections or topics which one feels to be too scantily treated (e.g., maturation and development). It is possible to find topics which are so concisely treated that they will probably need considerable supplementation before they can be adequately understood by the student. For example, the following constitutes the entire explanation of negative after-sensations: "The negative afterimage can best be understood by remembering that nerves are living tissue and that the chemistry of the receptors tends to remain in equilibrium. When stimulation pushes the photochemical balance in the eye off equilibrium, it tends to return, and the returning is a stimulus for the complementary colors" (p. 83). It is possible for the reader to find some scantily treated topics which he believes to be of greater importance than other topics which are treated at greater length. For example, six kinds of learning (motor skills, modification of perceptions, etc.) are discussed for 14 pages while efficiency of learning receives but three pages. A poem called "Similar Cases" is given a full page (p. 185), while the distribution of practice in learning receives a short paragraph of seven lines. Some instructors may doubt that enough experiments, especially those illustrating important points of methodology, have been presented in sufficient detail to give the beginning student the flavor of scientific research in psychology.

But in spite of the criticisms that may be made and the differences in opinion that may exist, the author has undoubtedly included much of importance and has presented many of the essentials of psychology in a concise way. He has not merely omitted difficult topics to make an easier and shorter text. Elementary statistical concepts and methods, for example, are presented in an appropriate place. The earlier chapters appear to be the most concise and would probably need the greatest amount of supplementation; the later chapters are longer and more completely developed. One may still feel, as does the reviewer, that the book might be better if it were longer and if more space had been used in discussing some topics. But for those who want a shorter text, *Essentials of Psychology* is well worth considering.

EDWARD E. ANDERSON.

Wilson College.

V. SCHILLER, PAUL. *Aufgabe der Psychologie: eine Geschichte ihrer Probleme.* Vienna: Springer-Verlag, 1948. Pp. 233.

The author, formerly of Budapest and recently a visiting lecturer at the University of Chicago, is now professor at the Bolyai University in Klausenburg, Romania. He gives a scholarly and authoritative survey of the theoretical positions of the major figures in the history of psychology from Plato through Tolman. Especially interesting is his discussion of contradictions which follow from Wundt's attempt to distinguish mediate and immediate experience and of Wundt's failure to follow his own theoretical program, in that he studied human functions and activities rather than conscious experiences. Some of the many other sections which deserve mention are those dealing with the theories of von Uexküll, American functionalism and behaviorism, and Brentano.

The book culminates in the author's *Handlungslehre*. The proper task of psychology is to study actions and their motivation. An action is defined as a specific behavior which is dependent on the total situation at a given time and which changes the situation; it includes the processes of arousal, steering, and execution united in a single whole. Conscious experiences are to be viewed as components of actions and not as belonging to an independent world of their own. The author regards his theories, which are expressed at a very abstract level of discourse, as representing a return to the holistic point of view characteristic of Aristotle and other thinkers who did not attempt to restrict psychology to the study of experiences or of external behavior and as a remedy for the tendencies which threaten to split modern psychology into many diverse and seemingly unrelated branches.

FRANK A. PATTIE.

University of Kentucky.

MACHOVER, KAREN. *Personality projection in the drawing of the human figure.* Springfield, Ill.: Charles C Thomas, 1949. Pp. ix+181. $3.50.

It would be unfortunate if we came to feel that our techniques in personality study were apart from our theories. If theories are valid, they should be demonstrable. If techniques are valid, it must be because they stem from some hypothesis which is basically true. This should hold for all aspects of our research even when practical considerations are most important, as it is understandable that they should be for the clinician whose primary concern is for the patient's welfare. It is, therefore, to be deplored that a prominent clinician, describing an important projective device, should state that: "Incentive for, and primary focus of, investigation centered around perfection of the drawing technique as a clinical tool for personality analysis, rather than around any theoretical hypotheses" (p. 20).

This statement is, of course, not adequately descriptive of the content of the book. The first part, "Personality Projection in the Drawing of the Human Figure," actually has a section on theoretical considerations. The second part, "Principles of Interpretation," draws heavily on psychoanalytic theory, as well as on empirical clinical experience. The third part, "Illustrative Case Studies," is an attempt to relate the theory of projection in drawing the human figure to examples of drawings accompanied by case histories.

The main purpose of this book, as stated by the author, is to begin to systematize the analysis of personality based on drawings of the human figure. Dr. Machover devotes several pages in the first section to a consideration of the sources of projection, which she labels the "psychic datum." She finds three major ones: common social meanings, the individual's own experience and symbol values. In discussing the common social meanings she writes: "It suffices to say that, from the empirical standpoint, such graphic communications occur regardless of age, skill or culture. One source to consider is the common social meanings that physical attributes tend to acquire in the course of social expression and intercourse" (p. 7). If the meanings are acquired as a consequence of social intercourse, how can they be free of culture, or even of age? Experience with the Rorschach, to which she often refers for corroborative evidence, indicates differences in interpretation for different cultures and for different age groups in our own.

In spite of such theoretical problems, the method seems to have considerable validity in clinical procedure. As one reads the illustrative cases, Dr. Machover's analyses can be followed without much argument. The question which often arises in connection with new projective techniques—"Would I see that if I didn't know the case history?"—is answered by her report on "blind" interpretations (pp. 25 to 27).

The drawing of the human figure is an extremely intriguing projective device. It is, admittedly, still in its infancy and communication of method is limited, as for most techniques which emphasize patterns of relationship rather than "scoring." It is very much to be hoped that readers will be stimulated by this book to further study and research. Dr. Machover introduces some challenging experimental problems. In spite of the format of the book, it is a progress report rather than a manual. It would be undesirable, with so many people working on this new method, to accept this report as a manual and mark the problem solved.

LILLIAN WALD KAY.

New York University.

MAYO, ELTON. *Some notes on the psychology of Pierre Janet.* Cambridge: Harvard Univ. Press, 1948. $2.50.

The author, in his preface, states that "this book is not intended for medical students or for those who have a special interest in the problems of psychiatry." It was written for "a few colleagues whose work demanded close attention to the difficult social, personal, and administrative problems of our time." While it is probably true that a knowledge of personality dynamics would be of service to such persons, it is hard to see in what way Mayo's presentation would be likely to give them valuable insights.

The book does not attempt a complete account of Janet's work, but his chief concepts are discussed sympathetically. Janet defines hysteria as a condition characterized by easy hypnotizability, suggestibility, dissociation, somnambulism, and shifting analgesia and paralyses. Most patients commonly diagnosed as hysterics Janet classifies as obsessives. It is as clear in reading Mayo's account as in reading Janet himself that many characteristics of his particular group of hysterics are the result of tutoring on the part of the psychiatrist. Mayo, in fact, says: "In several of these patients *Janet developed the capacity for separate response until it was possible to conduct two conversations with them simultaneously*" (p. 30).

Morton Prince, with Miss Beauchamp, demonstrated to what lengths this could be carried. It would be rare indeed were a layman to meet such a person in his daily life. Even the student of social administrative problems could manage quite well indefinitely without making use of this classification. Janet himself, according to Mayo, recognizes that "in the modern world the incidence of 'hysteria,' in the strict sense, has notably diminished" (p. 66). It would be difficult to defend Mayo's assertion that this change has occurred as a result of advances in nutrition and education rather than in improvement in diagnosis and therapy.

The obsessives, who include not only the people so classified ordinarily but some who would usually be considered to have hysteria or anxiety states (patients without somnambulism or divided personality), are the more important group according to Janet. Mayo believes them to be "the proper subject of psychopathological study in the strict sense" and the group most likely to concern a student of society. They may be divided into "frightened people" and "destroyers." They are unhappy, guilt-ridden, unable to relate themselves to others. Feeling that they live in an unreal world populated by phantoms, they are preoccupied with false alternatives, rituals, and fears which make it difficult for them to act or make decisions. It is important, says the author, for all students to understand the complexity of attention and the interrelation of action and reflection in order to detect the appearance of obsessive thinking in a social or industrial situation. The appearance of such symptoms will warn him that there is "some irksome constraint or feeling of insecurity imposed upon the individual or group . . . ." If he does not discover the source of this, the students will remain unable to understand the social situation he is studying.

While it is possible that, as Mayo says, "there is no real conflict between the observations of Janet and Freud. Indeed they work in different parts of the same field," this book suggests why Janet's work was relatively less fruitful. The rigidity of the classification and emphasis upon description rather than upon intervening "etiological" variables provide the clue.

University of Chicago. JAMES G. MILLER.

SYMONDS, P. M. *Dynamic psychology.* New York: Appleton-Century-Crofts, 1949. Pp. vii+413. $3.75.

This book is a revision, prepared for use as an undergraduate text, of the author's earlier book, *The Dynamics of Human Adjustment.* From the previous book, the following subject matter has been deleted: "(1) psychoanalytic interpretations which are least accessible to consciousness of most persons and hence which seem least convincing; (2) dynamics of abnormal states and conditions; and (3) unfamiliar vocabulary" (p. vi). The opening chapter is a more or less conventional introductory discussion of the relationship of psychology to other sciences, values to be derived from the study of psychology, etc. A new note is introduced, however, when the author says that "hundreds, indeed, thousands of reports of interviewing and analytic sessions have been reported and there is a repetition and consistency in them which go far toward overcoming the subjectivity of separate individual reports" (p. 9). While clinical data are rapidly attaining respectability, Symonds is probably among the first text-book writers to acknowledge explicitly its new status. The remaining chapters emerge from the author's essentially psychoanalytic frame of reference and deal primarily with the various "dynamisms" and the conditions in the environment and within the personality that govern adjustment.

The individual chapters of *Dynamic Psychology* are well organized and relevant points of view and findings are pulled together from a great number of sources. The reader is presented with what amounts to a review of the literature concerning the topic under consideration but reference is made to various writers in an expository style. The text, therefore, does not suffer from the dryness that might be expected in the ordinary "bibliographical" treatment of subject-matter. The over-all unity of the book could be improved. The author devotes only a paragraph to his general theme in the first chapter. In this paragraph he says, "This book deals with *dynamic psychology* which is interested in many states and processes, and particularly with those driving forces within the individual which seek their satisfaction in the world about" (p. 7). While this statement is satisfactory, it seems to the reviewer that the author might have done well to elaborate upon it toward the beginning of the book in his chapter on "The Ego and the Self." This chapter is the next to the last. Since it deals with the processes involved in the individual's relating himself to the environment there would be some advantage to bringing it forward and using it for purposes of developing a unifying theme.

Symonds has managed to write the kind of book that can be adapted to many purposes. The liberal use of illustrations and short cases makes it suitable for courses in adjustment that are primarily concerned with fostering the student's understanding of himself and others. The content can be counted upon to provoke discussion of various aspects of interpersonal relationships. The instructor who prefers a more formal presentation of theories of motivation, adjustment, and "personality" can also take his departure from the author's explanation of concepts, findings, and points of view of various writers. No attempt is made, however, to present a single systematic position. The instructor who is not content with an eclectic approach will not find the definitive framework he is looking for in this book, but he will find that it includes the ideas of a great many people.

FRANKLIN J. SHAW.

Purdue University.

THORPE, L. P., & KATZ, BARNEY. *The psychology of abnormal behavior.* New York: Ronald Press, 1948. Pp. xvi+877.

THORPE, L. P., & KATZ, BARNEY. *Workbook in the psychology of abnormal behavior.* New York: Ronald Press, 1948. Pp. x+124.

This encyclopaedic volume (877 pages) is intended for use as the textbook in the undergraduate introductory course in the psychology of abnormal behavior. The unusual range of topics covered, the logical and clear organization of the book as a whole and of each chapter, and the concise and direct style in which the book is written make it well suited for its purpose.

The major emphasis is placed on the dynamics of abnormal behaviors, in an effort to lead the student to an understanding of the mechanisms and adjustment processes utilized by maladjusted personalities in their efforts to maintain their psychological integrity. The authors appear to have achieved an unusually ecclectic coverage of the dynamic and etiological factors involved in abnormal behaviors by presenting the major competing views.

Part I (39 pages) is a brief but adequate introduction which reviews the social status and significance of abnormal behaviors, the history of research, points of view, and therapies, and the major contributions of foreign psychologists, psychiatrists, and others to the field.

Part II (115 pages) presents a general introduction to the dynamics of abnormal behaviors, criteria used for evaluating behaviors as abnormal, the nature of such behaviors, and a survey of organic and psychological etiological factors. Part III (58 pages) presents discussions of the nature and significance of symptoms, the classification of abnormal behaviors, and a review of abnormalities of the emotional, cognitive, and motor processes and of sleep. These two parts serve as brief but very adequate introduction to fundamental concepts regarding the dynamics, etiology, symptomatology, and evaluation of healthy and unhealthy behaviors. They also serve as adequate prerequisites to the parts which follow.

Parts IV through VII are concerned with specific adjustment problems and maladjusted personalities. The discussion of each disorder is presented in organized form in terms of incidence, symptomatology, etiology, treatment, prognosis, and prevention. This outline may be somewhat stereotyped but it provides the student with a definite study outline.

Part VIII (126 pages) reviews the field of therapies, including discussions of the objectives of therapy, methods of examination, and the major psychoanalytic, psychotherapeutic, and psychiatric methods. The presentation of these methods is primarily descriptive rather than evaluative or critical. The appendices present data on the classification of hospitals, statistics on the frequencies of patients with the various disorders, and a glossary of terms. The latter will be especially helpful to students. The book contains ample illustrative case reports. It is up to date on recent developments in diagnosis and treatment.

The authors have published a workbook for use with the text. It contains an objective test, suggested "research" projects, and discussion or "essay" questions for each chapter in the book. It also contains in table form chapter by chapter references in ten other books in the field of abnormal behaviors, so that students and instructors are provided with an easily used guide to assignments for collateral readings. Instructors will find the workbook a welcome source of materials for tests, of lists of relevant films and their sources, and of reading assignments.

WILBUR S. GREGORY.

The University of Redlands.

FRANK, L. K. *Projective methods.* Springfield, Ill.: Charles C Thomas, 1948. Pp. vii+86. $2.75.

BELL, J. E. *Projective techniques.* New York: Longmans, Green, 1948. Pp. xvi+533. $4.50.

The first book is a little volume which has been written by a man well known to psychologists and other students of human behavior as the director of the Zachry Institute of Human Development in New York City.

The first part of the monograph discusses the growing use of projective methods in psychology and psychiatry and shows how they are predicated upon a psychocultural conception of the emergence of the personality and its dynamic operation. The second part describes recent developments in scientific concepts and methods, especially in physics and chemistry and medicine. The third part reviews various approaches to the study and diagnosis of personality by way of preface to the more specific description of actual procedures. The fourth part outlines projective techniques and describes the five different varieties thus far developed, with interpretative comment upon the value and use of each variety, e.g., constitutive, constructive, interpretative, cathartic, refractive. The fifth part discusses reliability and validity of projective methods and the development of new criteria for assaying projective procedures.

Chapters one to three, approximately 41 pages, are a plea for newer types of psychological testing in light of changing concepts of personality theory. To bolster his position, the author borrows heavily, perhaps too frequently, from events in physics in particular and science in general. In subsequent chapters the author makes many statements without supplying the reference to the literature, which detracts from the book's usefulness to the advanced student. This limitation occurs rather frequently in the discussion of reliability and validity of projective methods.

Frank raises a tremendously important point which research psychologists as well as clinicians must some day attempt to rectify.

> While psychologists customarily rely upon the judgment of lay, naive individuals to validate their findings, it is difficult to find any scientific justification for such a procedure. If a scientific method or test is used, one cannot assume that the agreement or disagreement with its findings by a group of untrained, naive judges (as in matching) has any significance. Yet, this reliance

upon such judges is frequently cited as validating or invalidating a method (p. 66).

The appendix contains a partial bibliography of projective methods taken in part from Helen Sargent (pp. 68–86). This is a useful addition but has already appeared in the literature. This little volume is not prepared for the undergraduate major in psychology, nor will it be useful to the graduate student without preliminary training and experience in clinical psychodiagnostic methods. It has merit for theoretical exposition and for indicating the need for more research on projective tests before they earn their place in a scientific psychology.

While Frank's book is a theoretical treatment of projective methods, Bell's volume is also a useful survey of the varieties of techniques, and it attempts to serve more practical purposes. Thus, the author's aim is (1) to present a comprehensive review of the literature on projective techniques so that (2) research workers could apply batteries of projective tests to the same subjects for comparison of the respective merits. Thirdly, a description of each technique in one volume may permit its adoption as an introductory manual for the administration, recording, and interpretation of many of these instruments. It is agreed by all that a text for the training of graduate students in this area is already sorely needed. Fourthly, the volume is being presented in order to facilitate much needed research on these tools, especially in the study of validity and reliability. The final purpose of the author is to stimulate the invention of new methods.

Chapter I contains definitions of projection, general characteristics of projective techniques, some theoretical foundations, and sources of basic theorems underlying these methods. This last section refers to the contributions of Freud, Gestalt theory, clinical psychiatry, cultural anthropology, and learning theory and considers how all of these influences converge in a modern dynamic theory of personality.

The rest of the book is divided into five parts, as follows:

Part I discusses word association and related techniques, in addition to incomplete sentences, the tautophone, and story telling and completion procedures.

Part II consists of the visual stimulus techniques such as Rorschach, cloud pictures, thematic apperception test and its modifications, picture-frustration study and the Szondi test.

Part III is a treatment of expressive movement and related techniques and concerns the work of Allport and Vernon, Werner Wolff, the analysis of handwriting, Mira myokinetic psychodiagnosis, visual motor tests, drawing and painting, finger painting, completing pictures, the mosaic test, and analysis of voice and speech.

Part IV deals with play, drama, and related techniques.

Part V is an overview and evaluation of projective techniques, including the criteria for the adequacy of a projective technique, criticisms of projective techniques, and research needs for the future consideration of a more scientific approach to psychodynamic measurement.

Following each chapter the author presents a comprehensive and thorough bibliography. Several bibliographies contain references to the foreign literature as well as to the American journals. The most outstanding bibliography consists of 798 references and covers 50 pages (these are approximately all of the references to the Rorschach technique for the past 27 years since publication of *Psychodiagnostiks*). These features, as well as the author's lucid style of writing, recommend this volume as a text in clinical methodology for graduate students and as a reference book for psychologists and other clinicians concerned with human personality.

ARTHUR WEIDER.

University of Louisville School of Medicine.

WOLFF, WERNER. *Diagrams of the unconscious.* New York: Grune & Stratton, 1948. Pp. xiv+423. $8.00.

After a brief preface the book is divided into two parts containing seven chapters each. Part I gives a concise, though mostly non-critical, survey of the literature and introduces Wolff's ideas on how handwriting might be made to succumb eventually to experimental procedures. He believes he has discovered certain consistent "themes of personality" of the gestalt entity type which display the properties of form patterning plus sequential continuity. These themes gain their stability and properties by being anchored in the individual's neuro-muscular-physiological structure. As such, they are expressed in any *movements* of that structure. Handwriting, a product of some of those movements, when correctly analyzed depicts the individualistic characteristics of certain of those themes. Therefore, Wolff attempts to demonstrate by reports of experiments and analyses of illustrated signatures how writers employ the sub-properties of height, length, position and shape in organizing the form pattern and sequential continuity of their writing samples. The personality themes expressed are *dynamic symmetry* or proportional relationships, *rhythm* or periodicity of these characteristic properties, and *configuration* or the organization of the interrelationships of these different properties.

In my opinion: (1) his theoretical points are sufficiently sound to warrant investigation, (2) his individual signature analyses are instructive, but (3) his claims for experimental verification *cannot* be confirmed on the basis of the material included or referred to in this section. Wolff creates the unfortunate impression that he believes if he writes a statement with words of sufficient authoritative "ring," tosses in a few figures that have the appearance of acceptable size, the reader will accept statistical significance and substantiation of claimed relationships (pp. 58, 308). He uses eight pages (46–53), for example, in reporting one of his studies where, due to an inaccurate statistical procedure (failing to combine plus-minus scores), a non-existent significant trend was found; further, the problem the study was concerned with is only *slightly* related to the point under discussion. Incidentally, the graphs shown there are displaced.

Wolff refers his "lay reader" to Part II in preference to Part I, feeling it contains "the material demonstrating the principles that determine everybody's form and expression in his graphic" productions.

Here he all but ignores his efforts in Part I and falls back upon the classical approach to handwriting analysis. Throughout these seven chapters he implies that:

1. Most writing is composed of figures whose form gestalt is sufficiently similar to convention to be usually recognizable.
2. Written figures are characteristically changed by the writer from the printed norm but usually in non-distorting ways.
3. These changes are consistent for the individual.
4. The characteristic properties of these changes are accepted by more than one individual to be correlated with certain moods, degrees of intensity of response, attitudes, social roles and adjustment mechanisms.

Wolff uses many of his previously reported matching experiments plus other literature in substantiating these implications. The themes of personality are dropped in favor of a system of multi-traits. Chapter XIII tabularizes his various graphic elements, his grosser groupings of these, his list of personality traits with which he believes they are correlated, and a recording blank for systematizing and simplifying the analysis. The procedure is demonstrated on several famous signatures in the last chapter.

Here, too, Wolff desires to appear experimentally sound. However, (a) he subjectively selects his trait names and does not define them, (b) the great majority of these appear undefinable descriptively or operationally, (c) the experimental work reported lends little support to his claims of correlations between traits and graphic elements, (d) there are no tables of norms, and (e) his only stated caution is for the analyzer to be certain that an interpretation is done on the basis of the configurational organization of the graphic elements. How much better it would have been for Wolff to (1) frankly state that the work is still based on subjective clinical judgment, (2) that he still does not have experimental procedures which permit empirical evaluation of what he believes to be true from observations, and (3) that with these limitations he offers his findings to the *professional* reader, inviting assistance in solving the multitudinal problems involved. In that vein, this reviewer feels that Wolff will have more success by dropping his personality traits, returning to his concept of themes, and employing characteristic, describable response patterns with measurable properties.

This ought to be a worthwhile book. In many ways it is. Yet its obvious faults pervade so much of its material and influence, to the stage of inaccuracy, so many of its conclusions that one is hesitant to endorse it in spite of them. With adequate research many of its suggestions should lead to conclusive results. However, since many acceptable projective techniques are available, demonstration of considerable usefulness for this method is mandatory before others would be justified in spending the requisite research time and money.

FRED Y. BILLINGSLEA.

Tulane University.

HARRIS, NOEL G. (Ed.) *Modern trends in psychological medicine.* New York: Hoeber, 1948. Pp. xiv+450. $10.00.

The broad scope of investigation allowed by the title of these collected articles permits its authors to inquire into any crook and cranny of clinical psychology, psychiatry, child psychology, sociology, industrial psychology, mental hygiene and rehabilitation. The book may be said to have no beginning and no end; still, the editor in the introduction states that "an unorthodox layout may produce stimulus to thought not only in psychiatrists, but also in other members of our profession as well as in educationists and all those dealing with the character formation of children." Why a premeditated lack of organization or structure should be "a stimulus to thought" is not explained.

The range of this behavioral fantasia can best be understood by a list of the chapter titles in their order—The relation of psychological medicine to general medicine, the physiology of emotion, the importance of constitutional factors, the causative factors in mental disturbances, electrophysiology in psychiatry, diagnostic measures, marriage and the family, the future of child guidance, psychopathic personality, psychotherapy, recent technique of physical treatment and its results, further developments in abreaction, modern social and group therapy, principles of mental hygiene, character formation in relation to education, personnel selection, mental hygiene in industry, rehabilitation and the individual, psychological medicine and world affairs.

The first two chapters ("Psychosomatic medicine" and "The physiology of emotion") by N. G. Harris and S. Wright, respectively, present a brief but systematic survey of work that has been done in these fields, both here and in England, extending from approximately the early work of Cannon up to 1947. R. G. McInnes's chapter on "The causative factors in mental disturbances," consists of a well organized survey of the following factors—The genetic, constitutional, endocrine, psychological and social. F. Golla's article on electrophysiology in psychiatry and M. Jame's chapter on diagnostic measures have similar textbook value.

H. Stalker gives a sympathetic presentation of Henderson's theory concerning the nature of the psychopathic personality, but he concludes that the "key to the psychopathology of the psychopathic states is the concept of immaturity." This emotional immaturity is not explainable in terms of environment. Stalker rather suggests that it is "an inherent defect, meaning that the capacity to mature is to a greater or lesser degree wanting." This seems to be a sensible variation of Prichard's notion that the psychopath is a "moral imbecile."

The chapters on psychotherapy, recent technique of physical treatment, modern social and group therapy, personnel selection and mental hygiene in industry are all competent surveys of these fields.

The most interesting chapter in the book is Jules Masserman's discussion of psychological medicine and world affairs. His entire argument is based on his biodynamic principles of behavior, which include—the principles of motivation, environment interpretation, substitutive adaptation and neurotigenic conflict. Concerning the principle of motivation, he says "The behavior of all organisms . . . . is directed, in its deepest determinants towards the satisfaction of physiological needs." He leaves no room in his system for the existence of *acquired needs* which are not specifically physiological. But in 27 pages he attempts to explain national and internal problems and to suggest how most of them may be solved. His technique is unquestionably skillful. He refers to interesting analogies within the realm of his cat experiments in applying each of his four principles to the world at large. He, of course, protects himself each time, by reminding the reader that there is a difference between cats and men, but at the same time entices one to speculate concerning the importance of the existing similarities.

LIVINGSTON WELCH.

Hunter College.

BOOKS AND MATERIALS RECEIVED

ANASTASI, ANNE, AND FOLEY, JOHN P., JR. *Differential psychology.* (Rev. Ed.) New York: Macmillan, 1949. Pp. ix+894. $5.50.

BRITT, STEUART HENDERSON. *Social psychology of modern life.* (Rev. Ed.) New York: Rinehart, 1949. Pp. xvi+703. $4.50.

BURROW, TRIGANT. *The neurosis of man.* New York: Harcourt, Brace, 1949. Pp. xxvi+428. $7.50.

CHAPANIS, A., GARNER, W. R., AND MORGAN, C. T. *Applied experimental psychology.* New York: John Wiley, 1949. Pp. xi+434. $4.50.

CRONBACH, LEE J. *Essentials of psychological testing.* New York: Harper, 1949. Pp. xiii+460. $4.50.

DASHIELL. J. F. *Fundamentals of general psychology.* (3rd Ed.) New York: Houghton Mifflin, 1949. Pp. x+690. $4.00.

FLESCH, RUDOLF. *The art of readable writing.* New York: Harper, 1949. Pp. xiv+237. $3.00.

FRANK, JEROME. *Courts on trial.* Princeton: Princeton Univ. Press, 1949. Pp. xii+441. $5.00.

GOODENOUGH, FLORENCE L. *Mental testing.* New York: Rinehart, 1949. Pp. xi+609. $5.00.

HILLPERN, ELSE P., HILLPERN, ED. P., AND SPAULDING, I. A. *Bristow Rogers: American Negro.* New York: Hermitage House, 1949. Pp. 200. $3.00.

HOPPOCK, ROBERT. *Group guidance.* New York: McGraw-Hill, 1949. Pp. 393. $3.75.

KARPMAN, BEN. *Objective psychotherapy.* Monogr. Suppl. No. 6, *J. clin. Psychol.*, July 1948. Pp. 154.

LAWRENCE, MERLE. *Studies in human behavior.* Princeton: Princeton Univ. Press, 1949. Pp. x+184. $3.50.

LUNDHOLM, HELGE. *God's failure or man's folly?* Cambridge, Mass.: Sci-Art, 1949. Pp. 471. $6.75.

MEAD, MARGARET. *Male and female.* New York: William Morrow, 1949. Pp. 477. $5.00.

O'KELLY, LAWRENCE I. *Introduction to psychopathology.* New York: Prentice-Hall, 1949. Pp. xxi+733.

PICHOT, PIERRE. *Les tests mentaux en psychiatrie.* Paris: Presses Universitaries de France, 1949. Pp. 238.

POSTMAN, LEO, AND EGAN, JAMES P. *Experimental psychology.* New York: Harper, 1949. Pp. xiv+520. $4.50.

POTTER, MURIEL CATHERINE. *Reception of symbol orientation and early reading success.* Teach. Coll. Contr. Educ. No. 939. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949. Pp. viii+69. $2.10.

SALTER, ANDREW. *Conditioned reflex therapy.* New York: Creative Age Press, 1949. Pp. x+359. $3.75.

SARGENT, S. S., AND SMITH. M. W. (ED.) *Culture and personality.* New York: Viking Fund, 1949. Pp. 219. $1.50.

SHELDON, WILLIAM H. *Varieties of delinquent youth.* New York: Harper, 1949. Pp. xvii+888. $8.00.

SIMPSON, GEORGE G. *The meaning of evolution.* New Haven: Yale Univ. Press, 1949. Pp. xv+348. $3.75.

STONE, CALVIN P. *Case histories in abnormal psychology.* Stanford: Stanford Univ. Press, 1949. Pp. x+105. $1.75.

THORNDIKE, E. L. *Selected writings from a connectionist's psychology.* New York: Appleton-Century-Crofts, 1949. Pp. vii+370. $3.50.

VALENTINE, W. L., AND WICKENS, D. D. *Experimental foundations of general psychology.* (3rd Ed.) New York: Rinehart, 1949. Pp. xvii+472. $3.00.

VAUGHAN, ELIZABETH HEAD. *Community under stress.* Princeton: Princeton Univ. Press, 1949. Pp. 160. $2.50.

VERNON, P. E. *Personnel selection in the British Forces.* London: Little Paul's House, 1949. Pp. 324.

WALLIN, J. E. W. *Children with mental and physical handicaps.* New York: Prentice Hall, 1949. Pp. xxii+549. $5.00.

WEITZMAN, ELLIS, AND MCNAMARA, W. J. *Constructive classroom examinations.* Chicago: Science Research Associates, 1949. Pp. xvi+153.

WITTKOWER, ERIC. *A psychiatrist looks at tuberculosis.* London: National Ass. for the Prevention of Tuberculosis, 1949. Pp. 151.

WOLLNER, MARY HAYDEN. *Children's voluntary reading.* Teach. Coll. Contr. Educ. No. 944. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949. Pp. vii+117. $2.35.

YOUNG, KIMBALL. *Sociology.* (2nd Ed.) New York: American Book Co., 1949. Pp. viii+638. $5.00.

ZUCKER, CONRAD. *Psychologie des Aberglaubens.* New York: Grune & Stratton, 1949. Pp. 330. $5.50.

A critical review of the research on elementary school curriculum organization 1890-1949. Urbana: Univ. of Illinois Bulletin, 1949. Pp. 29.

Educacion. No. 7–8. Lima, Peru: Universidad Nacional de San Marcos, 1948. Pp. 261.

Guidance worker's preparation. A directory of the guidance offerings of colleges and universities. Misc. 3333. Washington, D. C.: Federal Security Agency, Office of Education, 1949. Pp. 45.

Vol. 47, No. 2 March, 1950

Psychological Bulletin

THE NATURE OF THE GALVANIC SKIN RESPONSE

ROBERT A. McCLEARY[1]

The Johns Hopkins University

The galvanic skin response or, more briefly, the GSR has already had a long history. It was discovered more than sixty years ago, and since then hundreds of papers have been written about it. In recent years, psychologists have been using it more and more as a tool in their experiments, and they shall probably continue to use it for many years to come. Finally, in view of the physiological and neuroanatomical evidence available on the GSR, its further study offers the investigator the possibility of a better understanding of the somatic mechanisms underlying behavior.

On the face of it, the GSR is a very useful measure in psychological experiments. It can be measured directly on a meter. It is a very sensitive response that will occur when many other responses fail. Like other autonomic responses, it usually cannot be voluntarily inhibited, when an adequate stimulus situation is present, though it may sometimes be voluntarily produced. It may occur in response to stimuli of which the individual is unaware—as has been demonstrated with hysterical patients (**56**) and post-hypnotic subjects (**48**). More recently, such "subconscious perception" has been demonstrated by the GSR method in non-hypnotized "normal" subjects (**50**, **51**). All these are distinct advantages that can be used in psychological experiments.

The GSR, however, has not lived up to expectations. For one thing, it now seems doubtful that the GSR represents a response that is connected to any specific category of mental activity. For another, the variability and the hypersensitivity of the response is often far too great to give the kind of results we need for good experiments. There is also the problem of the units of measurement—the proper scale—for the GSR (**43**). We cannot always be sure what measures of the GSR are significant and what are not. Furthermore, the GSR can be influenced

---

[1] The author is indebted to Dr. Clifford T. Morgan for his very generous help in the organization and writing of this paper.

Bureau Ednl. Psy. Re
DAVID HA

to a considerable degree by conscious set and the nature of experimental instructions (**9**). All these difficulties have tended to decrease the popularity of the GSR in psychological research.

There is still another basic question about the GSR. *What is it?* What are the physiological changes that give rise to it? That is a problem, which, if solved, might let us control somewhat better the factors that affect it in experiments. That is the problem for this paper. There have been a good many experiments on the problem, and several years ago Darrow (**23**) presented the conclusions that he could make from extensive original work. In this paper, the problem will be raised again, its history will be reviewed briefly, and some conclusions will be attempted.

A FEW WORDS ABOUT METHODOLOGY

This paper will not deal directly with problems of apparatus and technique. We must pay some attention to them, however, in order to get our terms straight. There is, moreover, hardly any field of endeavor where failing to understand such problems can so seriously affect both the results and the conclusions that one gets with the method. A few words about method are therefore in order.

Resistance-level and GSR. When a subject is hooked into a circuit for recording the GSR, there are two values one can measure. One is resistance-level. This is the static resistance that exists between the two electrodes of the circuit. It can be measured without any special stimulus. Introducing an "adequate" stimulus, however, such as a pistol shot or a disturbing word, causes a temporary change in the resistance-level, usually a lowering of resistance. The amount of this change is the second value that one can measure—the GSR.

It is now common knowledge that resistance level and GSR are related. In general other things being equal, the lower the resistance level, the smaller will be the GSR, and vice versa (**69**). Research workers sometimes measure both resistance level and GSR, but sometimes have used only one or the other. They have frequently made the assumption that any variable that lowers the level of static resistance is also operating in the transitory resistance drop which constitutes the GSR. This, of course, is not necessarily the case.

Exosomatic and endosomatic methods. Although we tend to think of the GSR as a change in resistance, this is only one of two types of changes that we can measure. When we measure resistance in a circuit, we usually apply a voltage across the resistance and measure current flow. Then, from Ohm's law, we can calculate the resistance. This method, when used with the GSR, has been called the *exosomatic method.* It is not necessary, however, to apply an external voltage. For most placements of the electrode on the skin, there is a resting electromotive force (voltage) between the electrodes. When some adequate stimulus is presented, there is some change, either in the EMF or the resistance or both, that registers on a galvanometer. We cannot, of course, know from the galvanometer reading just how much of the change is in resistance and how much in EMF (**28**). This method, however, will give a GSR. It has been named the *endosomatic method*.

Both methods have been used in experimental work. The endosomatic method, however, requires more sensitive recording techniques because of the low voltages afforded by the skin. For this reason, the exosomatic method has been more popular. In clinical use and in "lie-detector" work, it has been used almost exclusively.

There have been a few studies in which the results of the two methods have been compared. Wells and Forbes (**81**) made one such study. They concluded that the exosomatic method was more consistent and gave responses more closely related to affective states than did the endosomatic method. Another worker, Jeffress (**39**), has reviewed the literature and presented some research of his own on the problem. He was impressed with the high degree of relationship between the two methods—almost identical latent periods of reaction and a correlation coefficient of +.95 between amplitudes of deflection. Expressing agreement with these findings is Thouless (**73**). Finally, Darrow (**14**), using the two techniques alternately, could not find a one-to-one relationship between them, but showed that their respective results paralleled each other in a striking manner. In this report, therefore, the results of both methods will be used interchangeably without distinction.

There are many other interesting and important problems of technique in using the GSR. For example: the need for standardization of equipment, the importance of the type and placement of electrodes, and the relative virtues of A.C. and D.C. equipments. It is not the purpose of this paper to go into these problems. The interested reader may go to Darrow (**19**), Greenwald (**37**) and Forbes (**33**).

We have already referred to one other problem that has gotten some attention lately, namely, the unit of measurement for the GSR. This has been reviewed recently by Lacey *et al.* (**43**), who arrive at much the same conclusion as did Darrow (**21**) several years ago.

THE THREE VIEWPOINTS

Now we shall come to the main problem of the paper. What is the physiological basis of the GSR? Very early in the history of the GSR, three alternative notions were suggested, and these have had their champions down through the succeeding years. One is the *vascular theory*. It was expressed by Féré in 1888 who is commonly given credit—although there were at least seven papers before his (see **44**)—for first noting the electrical response of the skin. Certainly he first popularized it with psychologists. Féré thought that vasodilation was the physiological cause of the GSR. A second notion, that the GSR is the result of *secretory* activity of the sweat glands, was put forth by Tarchanoff in 1890. Finally, Sommer in 1902 first suggested that the change in electrical skin resistance was a matter of involuntary *muscular* activity. So within a period of 14 years from Féré's classical paper on the GSR, three schools of thought about its physiological basis had grown up.

There is now a considerable amount of evidence bearing on these three points of view. In 1929, Landis and Dewick (**44**), and later in 1932 Landis (**45**) brought out papers reviewing the literature on the GSR. Between them, these papers covered over 500 separate articles. Not all of these articles were concerned with the physiological nature of the GSR, but many of them were. This paper will not duplicate these two earlier reviews. It will, however, cite papers prior to 1932 that are especially relevant to the later work on the specific problem of whether the vascular, secretory or muscular theories of the GSR may be correct.

THE MUSCULAR THEORY

Sommer (1902) was mentioned just a moment ago. He was the first to suspect that muscular contraction might be the basis of the GSR. He felt that the reflex was simply a change in the contact between the skin and recording electrodes that came about because of involuntary muscular contractions.

This is a somewhat naïve view of the matter, and later investigators did not take it too seriously. They did, however, consider the possibility that the GSR might be the direct display of bio-electric changes in muscles. Sidis and Nelson (**71**), for example, suggested that the GSR was an EMF produced by covert muscular activity throughout the body. There are certainly some physiological facts that fit in with such a modified muscular theory. It is common knowledge, for example, that if some one cuts the motor nerve to an extremity, the GSR does not appear any longer (e.g. **7**). Then, too, as early workers pointed out, the GSR adapts or "fatigues" with repeated stimulation in a manner somewhat like the fatiguing of muscle fibers. And of course, as we all know, muscular activity does have its electrical counterpart. This kind of evidence, however, proves very little, for the motor nerves carry nerve fibers not only innervating striated muscles but also the blood vessels and sweat glands. Moreover, nearly all physiological systems will fatigue if stimulated repeatedly over a long enough period of time.

Still another kind of evidence has been cited by those in favor of a muscular theory. This is the effect on level of resistance of various tasks with clear muscular components (**83, 84**). Various maneuvers—squeezing a dynamometer, voluntary muscular contractions, sitting up, and so on—that increase muscular tension also decrease general skin resistance. It is not fair, however, to take such evidence as favoring the muscular theory, because all of these activities are vigorous enough to have many physiological repercussions including sweating and vascular changes. Furthermore, this is an effect on resistance level, not on GSR.

Worth mentioning in passing is one oddity reported over twenty years ago (1) and never thereafter confirmed or denied. A GSR was reported in a cat when electrodes were placed directly on an exposed muscle. The phenomenon was reported very sketchily and is hard to evaluate. We cannot be sure that the "GSR" was not related to the action potential of the muscle or to a change in the contact between the moving muscle and recording electrodes. The authors admit that the electrical change they observed was much smaller than the usual GSR.

All in all there is little impressive evidence for a muscular theory of the GSR. What evidence there is could just as well be interpreted to fit one of the other theories, the secretory or vascular theories. We shall come to these theories in a moment.

But even if the muscular theory is not the correct one, the relation of muscular effort to the GSR should not be neglected. If, for example, sweating is at the root of the GSR, the fact that muscular activity produces heat that must be dissipated from the body, ties muscular and secretory activity together. The relation between the two activities is therefore worth looking into (20, 29). With something like this in mind, one pair of investigators (32) presented subjects with a pedal-pushing task in which the tension in the pedaling system could be varied. As pedal-tension was increased, the palmar resistance fell. When action-potentials were led-off the muscles involved in the task, a similar relationship was found between the amplitude of action-potentials and the decline of skin resistance (30). The relationships, however, were not one to one. Incidentally, in this latter study, the EEG was also recorded, and there was a suggestion of a relationship between measures of skin resistance and action-potential on the one hand, and percent-time alpha rhythm.

Quite aside from any interest in the physiological basis of the GSR, French compared the GSR and records of finger-tremor by recording them simultaneously (34). Finger tremor, he found, has a markedly shorter latent period and a faster decay time than does the GSR. This fact fits in with the other evidence we have that the GSR is an autonomic response and is thus limited by the slower rate of conduction in the autonomic system than we find in the somatic motor system. French concluded that the GSR and finger-tremor have a different physiological basis—and probably a different psychological meaning.

THE VASCULAR AND SECRETORY THEORIES

Having disposed of the muscular theory, our serious attention should go to the vascular and secretory theories. These theories of the GSR are distinctly different, but the experimental evidence cannot be so neatly separated. Much of the evidence bearing on one theory also has something to do with the other one. It will be well, therefore, to consider them together.

Vascular theory. Both theories have been elaborated in the course of years and several varieties of each theory have been put forth.

Féré, when he originally proposed it, thought that *vasodilation* was the major event underlying the GSR. Later, however, as we came to know more about the physiological reactions involved in the startle-response and the GSR in startle, it began to look as though the vascular response, if indeed it is the basis of the GSR, might be *vasoconstriction.* Krogh (**42**), in fact, directly demonstrated in the ear-capillaries of rabbits a vasoconstrictor response to a startle-stimulus. Thus we find versions of the vascular theory put forth that provide for either constriction or dilation.

An explicit statement of this kind of vascular theory has been made by McDowall (**52**). He suggested two mechanisms for the GSR. (1) Whole blood, he pointed out, has a higher resistance than extracellular fluid. Consequently, the diminished blood content of the layers of the skin that comes about through vasoconstriction could cause a fall in skin resistance. (2) Also, because any change in the contour or tension of the skin can cause electrical changes—Denshams (**26**) had demonstrated this fact—McDowall suggested that vasodilation might have mechanical effects which would, in turn, cause a lowering of skin resistance. In this way, McDowall provided for both vasodilation and vasoconstriction in the GSR. By 1935, such a dual mechanism was necessary to hold on to a vascular theory, for Darrow had demonstrated that there are quite variable vasomotor responses connected with the GSR (**14**).

Secretory theory. This theory, as time went on, also came to rely on two mechanisms—in this case, alternative mechanisms. (1) One was the simple notion that the decreased resistance of the GSR comes about because of the presence in and on the skin of sweat—an electrolyte that would lower skin resistance (**55**). (2) The second suggested secretory mechanism was the more vague possibility that the GSR depends in some way on a change in the sweat-glands that precedes actual secretion—such, for example, as increased permeability of the cell-membranes whose resistance figures in the GSR circuit (**8, 73**).

The first possibility is rather easily discarded. For one thing, fluid-electrodes can be used to obtain the GSR, and one would expect these to lower the resistance so much that sweat could decrease it no more. For another, the time that it takes a GSR to subside, although it is slow, is not nearly as slow as it ought to be if it depended on the evaporation of sweat, which is a logical expectation from the theory. The second type of theory, however, is not so easily disposed of, partly because it is not so clearly stated and partly because experimental evidence on it is not so clear.

Is the Skin Necessary?

One way of deciding between the vascular and secretory theories is to see whether or not the skin is necessary for the GSR. The sweat-glands are found only in the skin, whereas blood-vessels, fortunately, are present everywhere in the body. If, therefore, the skin is not necessary to the GSR we could argue that the GSR is not a matter of sweat-gland activity. If the skin is necessary, we would still be left in a quandary between secretory and vascular theories.

You will recall a statement made above that the GSR was obtained with electrodes on the muscle of a cat (**1**). Unfortunately, however, we do not know whether to believe that it was really a GSR that was reported. There is only one other study that tests directly whether the skin is necessary. In this one (**26**) the outer skin layer—the epidermis—was taken off with a blistering agent. After that, *no GSR could be obtained*. It would be nice to have more corroboration of this finding, but it is all we have. From it we conclude that the *skin is necessary* to the GSR.

There are, fortunately, a number of data that fit in very well with the idea that the skin is necessary for the GSR. Some of them have to do with level of skin resistance rather than GSR. Richter (**60**) noted that there is a striking fall in resistance level when even a small needle prick is placed in the skin under the electrodes. This technique has sometimes been used purposely to cut out the resistance beneath one of the electrodes (the so-called "indifferent" electrode). Indeed we have studies (**26, 52**) suggesting that about 80 per cent of the total skin resistance is in the epidermis and the remaining 20 per cent is in the corium. Both of these studies, however, removed the epidermis and relegated the residual resistance to the corium. They neglected the fact that removing the epidermis would have injurious effects on the corium and might alter its usual resistance.

But even though there may be criticisms to make of any one study, the upshot of them all seems to be that both level of resistance and GSR are primarily properties of the skin. That takes us back where we started—we still cannot decide between secretory and vascular theories.

There is more evidence to come but this is a good place to bring in a rather different theory of the electrical nature of the skin—a hydration theory. Hemphill (**38**) has suggested this one. He believes that vasomotor or sweat-gland activity may account for the small resistance changes entailed in a GSR, but he believes that the mechanism of the static resistance is probably a different matter. He believes that the large swings that are seen in resistance level from day to day, occasionally amounting to 80,000 ohms or more, are too large to be a matter of secretory or vascular change. Certainly no clinical manifestations of vascular or secretory hyper- or hypoactivity are commonly evident. Conversely, even in acute cases of Reynaud's disease, when the hands are bone-dry, cold and all but avascular, there can be relatively small changes in resistance-level.

In an experiment with a small number of subjects, Hemphill has demonstrated the plausibility, although not necessarily the truth, of his idea. To get partial dehydration of the skin, he used brine baths at body temperature. These tended to elevate skin resistance. On the other hand, using corticotrophic hormone, which among other things aids in water-retention, he observed a somewhat lower skin resistance-level. These observations fit in with a point made earlier by Richter (60), who suggested that in non-palmar and non-plantar regions, the skin-resistance depends entirely on the condition of epidermal cells and responds mainly to local stimulation. Richter, however, did not attempt to show how the epidermal cells managed to change their level of resistance.

Whether or not there is anything to the notion that hydration changes level of resistance of the skin, it would be well worth knowing whether there are separate mechanisms, wholly or in part, for resistance level and GSR. As you saw earlier in the paper, there has been a tendency to throw observations of both phenomena into a common pool and treat them as though they have a common basis. Hemphill's suggestion is the only formal theory brought forward that postulates a physiological division between the two phenomena.

Now we can summarize what we know about the skin's role in resistance-level and GSR. The majority of the resistance we measure in the skin seems to be due to some physiological condition, rather than a simple mechanical condition, of the skin. So far as we know now, the skin is necessary for the resistance-change that we call the GSR. Whether resistance level and GSR have different mechanisms in the skin is a question that remains for solution.

Autonomic Drugs

With the multitude of drugs that we have to bring about autonomic reactions, one would expect that we could use autonomic drugs very nicely to get at the physiological basis of the GSR. That, however, has not proved to be the case, and in a moment we shall see why.

One trouble is that sweat-glands have a unique position in the autonomic system. Although they get their innervation from the sympathetic (thoraco-lumbar) system, they behave pharmacologically as 'cholinergic' end-organs (**11, 12**). This essential fact was not recognized in work with the GSR until 1936 (**24**).

Furthermore, investigators have often assumed that drugs that lowered resistance level and consequently diminished the GSR were having some direct effect on the GSR. But you will recall that changes of resistance level from any cause generally affect the amplitude of the GSR. Since autonomic drugs as a rule affect both level and GSR, this is an obvious point of confusion and interpretation of results is difficult indeed.

Another trouble is that most of the autonomic drugs are not highly specific. Their actions are not strictly limited to one or the other autonomic end organ. Atropine, for example, both decreases secretion and produces vasodilation. Pilocarpine increases secretion and produces vasodilation. Adrenalin is a vasoconstrictor (even reported to be a vasodilator in weak concentration) and an inhibitor of sweat-gland activity. Pilocarpine, moreover, exhibits nicotinic action at the ganglia but the antagonistic muscarinic effect at peripheral end organs. To this confusing state of affairs, add the fact that the vascular theory of the GSR is prepared to accept either vasodilation or vasoconstriction, and it is not surprising that results from this direction have been highly equivocal.

But even though we can not expect too much, let us look briefly at some of the experimental findings with autonomic drugs. With the exception of Waller (**81**) investigators agree that atropine elevates the resistance level (**1, 53, 57**). The effect of atropine on the GSR is less clear, however. From two sources (**44, 81**) we are advised that it lowers the GSR, while others (**53, 80**) feel that it has no effect on the GSR. Pilocarpine lowers the resistance level—on that there is unanimous agreement (**1, 53, 57**). However, only one investigator, Fauville (see **44**), reports on its effect on the GSR, and he says that it lowers the GSR.

Because the vascular effects of atropine and pilocarpine are thought to be similar, we ought to get some clues from these experiments. The fact that atropine raises resistance-level and pilocarpine lowers it ought to mean that secretory changes—or at least non-vascular changes—account for resistance level. It is unfortunate that the data on the effects of the drug on GSR are either inconsistent or scanty, otherwise we might make similar inferences about the GSR.

Adrenalin and its effects on skin resistance have also come under scrutiny, but the picture is not too clear. In the early work, there was a disagreement between Aveling and McDowall (see **1**), but two more recent studies, both having the benefit of improvements in technique, agree that adrenalin increases resistance-level (**52, 54**). As to the GSR Darrow demonstrated that it is diminished after intravenous injection of adrenalin (**23**). Because adrenalin in moderate doses is a cutaneous vasoconstrictor, it seems more likely that its effect on resistance-level is a matter of inhibiting sweat-gland activity. Whether this is also the basis of the reduced GSR, however, remains an open question.

Clinical Impressions

There is a vast literature on the characteristics of the GSR in various neuropsychiatric syndomes, but it would be beside the point to take it up here. There are, however, a few clinical reports that have direct implications for the physiological nature of the GSR, and we shall consider them briefly here.

One of the reports (82) is that GSR's are difficult to obtain from patients with arteriosclerosis ("hardening of the arteries"). This fact does not necessarily weigh heavily in favor of a vascular theory, however for patients in the older age-groups where arteriosclerosis is most common are also at the age at which senile hypohydrosis (decreased perspiration) is commonly making its appearance.

Somewhat more informative are three reports of the electrical characteristics of patients that had pathological conditions of the sweat-glands. Without giving the original reference, McDowall (52) reported that Gilchrist was able to get a GSR in a patient who lacked sweat-glands. The absence of sweat-glands was histologically confirmed. If we could believe this report fully, and particularly that there were no vestigial sweat-gland elements present in the patient, we would have conclusive proof against the secretory theory. Richter (60) reported on a similar type patient, with scleroderma. The patient's resistance level did not respond in the usual way to atropine. With regard to the GSR, the patient gave an extremely atypical response of long latency and very slow decay. Golla (36) finally, examined a patient with hyperhydrosis (excessive sweating) and reported the GSR to be normal. Some think this fact denies a sweat-gland theory, but that is not necessarily the case, for it is reasonable that even a hyperactive gland might increase its activity with stimulation.

Anatomical Considerations

In arguing the pros and cons of sweat-gland and blood-vessel theories, reviewers and investigators have appealed to various sorts of anatomical facts. One of these is the distribution of blood-vessels and sweat-glands over different parts of the surface of the body. It is well-known that the largest GSR's are obtained in the palmar and plantar areas where the concentration of sweat-glands is the greatest. Yet these same areas are also richly endowed with blood-vessels (52). We, therefore, get little help from this quarter.

Sommer (see 52), however, reports somewhat larger GSR's from the finger tips than from palmar areas, and we know that the capillary plexi in the *corium papillae*—the vascular structure nearest the avascular epidermis—are very concentrated on the palms and soles and are increasingly dense near the tips of the digits (4). On the other hand, Sommer, (see 52) states, perhaps without too much evidence, that sweat-glands are less prevalent in the finger tips than on the palms. If we put all of these points together, we should be inclined to favor a vascular theory of the GSR.

Our best information about the resistance level of different parts of the body comes from a series of researches by Richter. In this series his principal concern has been the differences in resistance levels of various bodily areas—mainly palmar vs. dorsal hand areas. He recorded separately and simultaneously from the two areas of the hands while manipulating such variables as heat, emotional stimuli and sleep. From his studies he has been able to show that resistance-changes in the two areas are often different in both amplitude and direction (**60**). The palm, according to Richter, responds in one-to-one relationship with sweat-gland activity, but such is not always the case with the dorsum. To him the difference in the behavior of the two areas can be understood by assuming that the epithelial cells of the dorsum are more responsive to purely "local" stimulation. Notwithstanding this notion, however, the dorsum did respond to atropine and to a warm bath in just the way we would predict if sweat-glands were the basis of the response.

This same article also describes the effect of denervating a monkey's arm. The palmar area, after an appropriate postoperative period of recovery, maintains an unusually high resistance level and does not give any GSR. The dorsum, on the other hand, shows very little post-operative change. Thus, it seems clear that there is some important difference between palmar and dorsal areas. What the difference means, however, we do not know yet.

One final anatomical finding remains to be noted. As mentioned above there has been considerable interest in the patterns of low resistance occurring normally over the body's surface (**58**, **65**). These studies have had considerable neurological value. One of the naturally occurring low resistance areas is over the palm, extending up over the dorsal tips of the fingers—the area where GSR's are reportedly the largest. This region, it might be mentioned in passing, is also the skin-area in which, histologically, no sebaceous glands can be found (**4**).

Vascular Considerations

Early investigators employed a sphygmomonometer (blood-pressure cuff) in an attempt to alter directly the vasomotor reactivity in an extremity and then to observe the effect of this procedure on the GSR. Their findings were equivocal (**27**). We can ignore them, however, for we know now that a tight blood-pressure cuff does *not* completely paralyze the vascular responses in an extremity (**42**). Furthermore, the extreme vascular congestion that results from such a procedure might well be expected to disrupt to some extent all the physiological systems in an extremity.

Darrow (**14–25**) has given us a series of studies over a period of years that contribute materially to our understanding of the mechanism of the GSR. In one of these studies (**14**), he simultaneously recorded GSR, plethysmograph changes, blood-pressure responses and, by an ingenious apparatus, the moisture changes on the surface of the skin. Disregarding the changes in volume of the extremity that were occurring subsequent to blood-pressure changes (a precaution not taken in some of the earlier plethysmographic studies), he showed quite strikingly that changes in volume, which could be attributed to vasomotor responses, were not related in any systematic way to the occurrence or amplitude of the GSR's. In a later paper when plethysmograph and electrode were arranged to record from the same finger, he confirmed these observations (**17**).

Darrow, in fact, has given good reason to believe that blood-pressure responses and GSR's have a different psychological significance. He suggests that increased blood-pressure tends to occur after stimuli of disturbing "ideational" content, while GSR's occur more frequently in the face of pure "sensory" stimuli—usually of a startling nature.

Opposed to the notion of complete independence of GSR and vascular phenomena, however, is another study by Lauer and Smith (**47**) in which GSR and pulse-rate were recorded simultaneously and "time for maximal deflection" was the measure of the GSR. In this case, there was a correlation coefficient of .80 between pulse-rate and GSR.

It would seem unlikely that vasomotor responses are in any causal way related to the GSR, but it is clear that cardio-vascular changes do occur in response to various psychic stimuli. The nature of these cardio-vascular responses and what constitutes their adequate stimuli is, of course, not within the scope of this paper.

All of these results can make sense, however. There may well be a high correlation between vascular and secretory changes in terms of some of the stimulus conditions that arouse them. There may not be—in fact, Darrow's experiments made it pretty clear that there is not—any causal relation between the vascular responses and the GSR. If this evidence and reasoning be correct, we finally have grounds for believing that some aspect of sweat-gland activity underlies the GSR.

Considerations of Secretory Activity

Darrow's work (**14–25**) lent heavy support to the possibility that presecretory activity of the sweat-glands forms the basis of the GSR. In his original paper mentioned above (1927), he demonstrated clearly that the electrical changes in the skin preceded quite consistently, by about one second, the increase of moisture on the surface of the skin as measured by his equipment. Later (**18**) he published enlarged microfilms of a palmar skin area on which a shadow of a galvanometer needle was superimposed. In this way he was able to correlate galvanometer deflection and the appearance of droplets of perspiration on the skin surface. In this same study, he showed that, at high resistance levels, large GSR's would occur with minimal perspiration, whereas at low resistance levels, minimal GSR's would occur even with considerable perspiration. This bears an obvious relationship to the fact that amplitude of GSR is a function of the resistance-level. On this basis, Darrow showed that change in electrical conductance more closely mirrored the secretory response than did change in electrical resistance (**21**). He suggested that conductance would be a more meaningful measure of the response.

Carrying this approach farther, Freeman and Darrow (**23, 29**) attempted to show a relationship between resistance change and weight loss from insensible perspiration (using a modification of the Sauter balance for weighing). They got a trend toward positive correlation, but it was neither consistent nor high. To explain their results, they pointed out that insensible perspiration is present on many surfaces, while they were recording electrically from only a few surfaces at a time. Another possibility, too, is that the results were contaminated by decreases in the water lost in breathing. We know that respiration is slowed in startle responses, yet GSR's are produced in startle responses. Thus, the decreased evaporation in respiratory passages might offset increases in perspiration of the skin that go along with the GSR.

Tying all the pros and cons together, it certainly seems likely that the GSR is dependent on sweat-gland activity. More specifically, the GSR would appear to be the result of some pre-secretory change in the sweat-glands.

ADAPTIVE VALUE OF THE GSR

Assuming this to be the case, why should this sweat-gland activity take place in response to various kinds of sudden or threatening stimuli? In several later papers (**22, 23, 24, 25**), Darrow suggested one explanation. He postulated that it is a preparatory and facilitative reflex to assist in grasping responses and to increase tactile sensitivity in the face of a threatening situation. He presents interesting evidence for this, but it will not be covered in this paper.

An alternative explanation of this apparent relationship between GSR's and perspiration entails the consideration of temperature regulation as one aspect of the response. Since the organism's general metabolism accelerates when presented with some 'startling' stimulus situation, it might be reasoned that the sweat-gland activity is simply one result of the body's attempt to assure adequate heat loss in the face of the stepped-up metabolism. It might then be regarded as a transitory homeostatic mechanism or as the "over-shooting" of a bodily mechanism called precipitously into action.

There is some experimental work bearing on the relationship between GSR, metabolism and temperature loss. Freeman and Giffin (**31**) studied the relationship between palmar resistance level and BMR (basal metabolic rate) in an attempt to determine whether there was a close enough relation to suggest the dependency of resistance level on the need for temperature loss. They interpreted their results as denying such a relationship. In a series of studies concerned with the effect of sleep on the resistance level (**57, 58, 59, 62**), Richter has shown a differential effect of sleep on resistance depending on whether he was recording from palmar or non-palmar areas. Palmar resistance tends to rise, while non-palmar falls. The author felt that this mirrored the fact that during normal sleep one's palms are normally dry while—temperature regulation requiring it—the necessary sweating takes place in non-palmar, non-plantar areas. In sleep, then, the palms do not seem intimately concerned with temperature control.

How the GSR comes to be included in the organism's repertoire of reflexes must remain for now a matter for speculation.

NERVOUS CONTROL OF THE GSR

There remains to be considered, briefly, the central and peripheral nervous connections of the GSR. There is of course a voluminous literature on the nervous correlates of autonomic phenomena in general. However, in the section below, we shall limit our inquiry only to those researches bearing the most directly on our problem.

Cortical localization. Since the GSR can be affected to a considerable degree by conscious processes, it seems likely that there is a cortical center or centers mediating control over the response. Furthermore, a review of the work to date reveals excellent agreement on where this center is.

Early in this century, German workers (Bechterew and Winkler) directed attention to the pre-motor cortex as being involved in the GSR (see **25**). Numerous researches since then (**46, 68, 72, 77, 79**) have confirmed this observation through the use of cortical stimulation technique. One of these studies (**46**) also reported eliciting a GSR by electrical stimulation more caudally in the cortex (posterior to the hind-limb motor representation) but this report remains to be confirmed.

It is interesting that the pre-motor cortex (*Brodman 6*) should turn out to be important in the GSR, because completely independent work reveals that this area is of importance in the sweat response to high environmental temperature. In a series of monkeys with unilateral extirpation of area *6*, it was shown that these animals had markedly diminished sweat responses on the side contralateral to the cortical lesion (**41**). The same author examined neurological patients with various non-specific lesions of area *6* (the so-called pre-motor syndrome) and was able to demonstrate both vasomotor and secretory losses on the contralateral side (40).

There is another point about patients with the pre-motor syndrome that might be mentioned in passing. The point concerns Darrow's suggestion that the secretory response is a reflex to aid in grasping. It so turns out that patients with the pre-motor syndrome do often show actual motor deficiencies in grasping responses. This same type of motor loss has been reported in monkeys following extirpation of the pre-motor area (63). Moreover, with a subject lying on his side, larger GSR's can be recorded from the uppermost hand than from the lowermost hand (22). Similarly, patients with the pre-motor syndrome, when placed in a like position, demonstrate more pronounced forced grasping reflexes in the uppermost hand. Regardless of the applicability of these clinical observations, there appears to be enought evidence available to say that Brodman's area *6* is one of the main cortical centers for the GSR.

Subcortical center and tracts. There are several reasons for thinking that there are also sub-cortical centers capable of modifying the GSR. For one thing, as you will recall from early in this paper, subjects can give GSR's to stimuli of which they are unaware. Furthermore, following pre-motor area lesions (40) and even after certain spinal cord lesions (10), the GSR remains intact and even hyperactive. The story on these sub-cortical centers, however, is not a clear one, as we shall see.

In view of the numerous autonomic functions which appear to have some representation in the hypothalamus, it was not out of line with expectation when it was shown that electrical stimulation of the tuber-cinereum of cats produced a GSR in all foot-pads (75). It was shown several years later, however, that electrical stimulation of the pre-motor cortex could produce a GSR even with the tuber-cinereum excised (77). This observation does not imply that there is necessarily a sub-cortical "center" other than the tuber-cinereum. But it makes it necessary to find some alternative sub-cortical efferent pathway that does not include the tuber-cinereum.

There are several things to be said about these sub-cortical pathways. It has been shown that after transecting the cerebral peduncles, stimulation of the pre-motor area will no longer produce a GSR (77). This observation suggested that the pyramidal tracts were the cortico-spinal pathway for the GSR. On the other hand, later work showed that sectioning the peduncles more caudally did not disrupt the GSR pathway, and this casts serious doubt on the importance of the pyramidal tracts in this response. It certainly makes more sense if the pyramidal tracts are not involved, since stimulation of the motor area does not produce a GSR (46). Langworthy and Richter, in consideration of the neuroanatomical evidence available, suggested that the fronto-pontile and temporo-pontile tracts were good possibilities, and in partial confirmation of this point showed that GSR's could be elicited by stimulation of certain extra-pyramidal areas in the sub-cortex.

Spinal tracts. Several attempts to delineate the spinal tracts that carry the efferent limb of the GSR have been made. Poor agreement among the studies, however, makes difficult any clear conclusions. For example, there are two studies (**77**, **79**) pointing out the combined crossed and uncrossed course of the spinal tracts for the GSR. The responses to electrical cortical stimulation were, however, noted to be consistently stronger on the *ipsilateral side*. These authors, incidently, were the investigators who thought the pyramidal tracts were involved in the response and they postulated that the uncrossed pyramidal fibers were autonomic in character. They also demonstrated what was thought to be a separate spinal tract from the hypothalamic center.

In contrast to the above findings is a more recent study employing extirpation technique (**68**). Following removal of area *6*, no effect whatsoever on the GSR of the ipsilateral side was noted, while there was complete absence of "psychic" GSR's on the *contralateral side*. GSR's to physical nociceptive stimuli, however, could still be obtained on the contralateral side, as long as the stimulus was applied in the same segment from which the GSR was being recorded. The author concluded that area *6* was necessary for "psychic" GSR's and for segment to segment transmission of the afferent end of the reflex. Unfortunately, this study did not report the results of recording from any segment other than the forelimbs. The important point for us to note is that he concluded that the spinal tracts involved were completely *crossed*.

If we are willing to assume that the GSR does in fact depend in some way on sweat-gland activity, we might expect some help from studies directed to the question of the spinal course of sweating responses. In four studies (**4**, **5**, **40**, **41**), the conclusion is offered that the nervous pathway for secretory responses is completely *crossed*. In view of the reasonable doubt remaining as to the causal importance of secretory responses in the GSR, further experimentation aimed at determining whether the GSR is crossed or uncrossed in the spinal cord would be of considerable interest.

Peripheral nervous mechanisms. As is often the case, when we come to the peripheral nervous system, the picture is much more clear. Numerous investigations (e.g. **54**, **66**, **74**, **77**) have shown by direct surgical approach that the GSR is dependent on an intact sympathetic chain, and conversely, that direct stimulation of the chain or the post-ganglionic fibers involved will produce a GSR. More quantitatively, the amplitude of the GSR is a direct function of the intensity of such sympathetic stimulation and the number of post-ganglionic fibers left intact (**78**). Moreover, the conduction rate in the post-ganglionic fibers has been determined. The rate is an order of magnitude that fits in nicely with the latent period, duration, and tetanization rate of the GSR (66). Finally, the GSR can be obtained in all extremities even after mid-thoracic spinal transection in cats (61). (This latter type of preparation, incidentally, might offer interesting possibilites for the study of spinal conditioning—using the GSR as the CR.) Clearly, the sympathetic chain and its post-ganglionic out-flow provide the final-common-path for the GSR.

Several studies suggest a surprising degree of autonomy on the part of the sympathetic chain. For example, the skin resistance elevation commonly observed after severing the peripheral motor nerve (7) is also observed if the sympathetic-fiber action is done at the level of the pre-ganglionic fibers involved—but is considerably less (74). Further, it has been suggested (67) that the sympathetic ganglia may mediate GSR's independently of the CNS.

SUMMARY

With a 60-year history behind it, we still find the GSR being used in psychological research. In all respects, however, it has not lived up to what was originally expected of it.

Research workers have spent a lot of time trying to work out the physiological basis of the GSR. Throughout this work, three main kinds of ideas about the important physiological factor keep coming up: (1) muscular activity, (2) vascular activity and (3) sweat-gland activity. We can be fairly certain now that it is some aspect of sweat-gland activity. We still do not know precisely how the sweat-glands manage to produce a GSR. Actual sweating goes along with the GSR, but the electrical change appears to occur before the secretion.

The neuro-anatomical pathways of the GSR have come in for their share of investigation. We can be sure that the sympathetic-chain and its post-ganglionic fibers are the final common path of the reflex. Workers also agree on the importance of the pre-motor cortex (*Brodmann 6*). The sub-cortical pathways, however, have caused more trouble. The tuber-cinereum apparently plays a role, but there is probably also some other efferent route in the sub-cortex. Whether the spinal tracts are crossed or uncrossed has not been settled. Physiological work on the "sweating-response" would make you think that the spinal tracts for the GSR would be crossed. Several studies point to a considerable degree of autonomy at the level of the spinal-cord and sympathetic-chain.

BIBLIOGRAPHY

1. Aveling, F., & McDowall, R. J. The effect of the circulation on the electrical resistance of the skin. *J. Physiol.*, 1925, **60**, 316–321.
2. Aveling, F., & McDowall, R. J. A note on the psychogalvanic reactions of anesthetised cats. *Brit. J. Psychol.*, 1925, **16**, 50–52.
3. Bayliss, W. M. The electrical changes in electrical tissues. *Brit. med. J.*, 1919, **1**, 670–671.
4. Beaton, L. E., *et al.* Thermoregulatory pathways in the cat's brain stem. *J. Neurophysiol.*, 1943, **6**, 29–35.
5. Beaton, L. E., *et al.* Spinal distribution of thermoregulatory pathways in the monkey. *J. Neurophysiol.*, 1943, **6**, 37–38.
6. Brickner, R. M. Certain characteristics of cortical influence over the sympathetic nervous system in man. *J. nerv. ment. Dis.*, 1930, **71**, 689–713.
7. Bruesch, S. R., & Richter, C. P. Cutaneous distribution of peripheral nerves as determined by the electrical skin resistance method. *Johns Hopkins Hosp. Bull.*, 1946, **78**, 235–260.
8. Cummins, H. The skin and mammary glands. From H. Morris, *Human anatomy.* (10th Ed.) Philadelphia: Blakiston, 1942.
9. Cook, S. W., & Harris, R. E. The verbal conditioning of the galvanic skin reflex. *J. exp. Psychol.*, 1937, **21**, 202–210.
10. Craig, C. B., & Hare, C. C. Sweating reactions in patients with diseases of the spinal cord. *Arch. Neurol. Psychiat.*, 1935, **33**, 478–491.
11. Dale, H. H. Nomenclature of fibers in the autonomic system and their effects. *J. Physiol.*, 1933, **80**, 10p–11p.
12. Dale, H. H., & Feldberg, W. The chemical transmitter of nervous stimuli to the sweat glands of the cat. *J. Physiol.*, 1934, **81**, 40p–41p.
13. Darling, R., & Darrow, C. W. Determining activity of the autonomic nervous system from measurements of autonomic change. *J. Psychol.*, 1938, **5**, 85–89.
14. Darrow, C. W. Sensory, secretory and electrical changes in the skin following bodily excitation. *J. exp. Psychol.*, 1927, **10**, 197–226.
15. Darrow, C. W. Electrical and circulatory responses to brief sensory and ideational stimuli. *J. exp. Psychol.*, 1929, **12**, 267–300.
16. Darrow, C. W. Differences in the physiological reactions to sensory and ideational stimuli. *Psychol. Bull.*, 1929, **26**, 185–201.
17. Darrow, C. W. The galvanic skin reflex and finger volume changes. *Amer. J. Physiol.*, 1929, **88**, 219–229.
18. Darrow, C. W. The relation of the galvanic skin reflex recovery curve to reactivity, resistance level and perspiration. *J. gen. Psychol.*, 1932, **7**, 261–273.
19. Darrow, C. W. Considerations for evaluating the galvanic skin reflex. *Amer. J. Psychiat.*, 1933, **13**, 285–298.
20. Darrow, C. W. Quantitative records of cutaneous secretory reactions. *J. gen. Psychol.*, 1934, **11**, 445–453.
21. Darrow, C. W. The significance of the galvanic skin reflex in the light of its relation to quantitative measurements of perspiration. *Psychol. Bull.*, 1934, **31**, 697–698.
22. Darrow, C. W. Palmar skin conductance (sweating), an accessory to activity of the postural mechanisms. *Psychol. Bull.*, 1935, **32**, 720–721.
23. Darrow, C. W. The galvanic skin reflex (sweating) and blood pressure as preparatory and facilitative

functions. *Psychol. Bull.*, 1936, **33**, 73–94.

24. Darrow, C. W. The palmar galvanic skin reflex (sweating) and parasympathetic activity. *Amer. J. Psychol.*, 1936, **48**, 522–524.
25. Darrow, C. W. Neural mechanisms controlling the palmar galvanic skin reflex and palmar sweating. *Arch. Neurol. Psychiat.*, 1937, **37**, 641–663.
26. Densham, H. B., & Wells, H. M. The mechanism by which the electrical resistance of the skin is altered. *Quart. J. exp. Physiol.*, 1927, **18**, 175–184.
27. Densham, H. B., & Wells, H. M. The effect of the circulation on the skin-constrictor (psychogalvanic) reflex. *Quart. J. exp. Physiol.*, 1927, **18**, 283–289.
28. Dorcus, R. M., & Harrell, T. W. Some relations between initial body polarity and polarity during the psychogalvanic reflex. *J. gen. Psychol.*, 1937, **16**, 489–491.
29. Freeman, G. L., & Darrow, C. W. Insensible perspiration and the galvanic skin reflex. *Amer. J. Physiol.*, 1935, **111**, 55–63.
30. Freeman, G. L. Cortical autonomous rhythms and the excitation levels of other bodily tissues. *J. exp. Psychol.*, 1940, **27**, 160–171.
31. Freeman, G. L., & Giffin, L. L. The measurement of general reactivity under basal conditions. *J. gen. Psychol.*, 1939, **21**, 63–72.
32. Freeman, G. L., & Simpson, R. M. The effect of experimentally induced muscular tension upon palmar skin resistance. *J. gen. Psychol.*, 1938, **18**, 319–326.
33. Forbes, A. The limiting AC frequency for the exhibition of the galvanic skin (psychogalvanic) response. *J. gen. Psychol.*, 1935, **13**, 188–193.
34. French, J. W. A comparison of finger tremor with the galvanic skin reflex and pulse. *J. exp. Psychol.*, 1944, **34**, 494–505.
35. Fulton, J. F., *et al.* Autonomic representation in the cerebral cortex. *Amer. J. Physiol.*, 1934, **109**, 37.
36. Golla, F. L. The objective study of neurosis. *Lancet*, 1921, **2**, 215–221.
37. Greenwald, D. V. Circuits now available for the measurement of electrodermal responses. *Psychol. Bull.*, 1935, **32**, 779–791.
38. Hemphill, R. E. Electrical resistance of the skin. *J. ment. Sci.*, 1942, **88**, 285–305.
39. Jeffress, L. A. Galvanic phenomena of the skin. *J. exp. Psychol.*, 1928, **11**, 130–144.
40. Kennard, M. A., *et al.* The syndrome of the premotor cortex in man. *Brain*, 1934, **57**, 69–84.
41. Kennard, M. A. Vasomotor disturbances resulting from cortical lesions. *Arch. Neurol. Psychiat.*, 1935, **33**, 537–545.
42. Krogh, A. *Anatomy and physiology of the capillaries.* New Haven, Conn.: Yale Univ. Press, 1929.
43. Lacey, O. L., *et al.* An analysis of the unit of measurement of the galvanic skin response. *J. exp. Psychol.*, 1949, **39**, 122–127.
44. Landis, C., & Dewick, N. The electrical phenomena of the skin (psychogalvanic reflex). *Psychol. Bull.*, 1929, **26**, 64–119.
45. Landis, C. The electrical phenomena of the skin (psychogalvanic skin response). *Psychol. Bull.*, 1932, **29**, 693–752.
46. Langworthy, O. R., & Richter, C. P. The influence of efferent cerebral pathways upon the sympathetic nervous system. *Brain*, 1930, **53**, 178–193.
47. Lauer, A. E., & Smith, E. A. A quantitative study of the relation between pulse and breathing changes and electro-biochemical responses. *Amer. J. Psychol.*, 1932, **44**, 732–739.

48. Levine, M. Psychogalvanic reflex to painful stimuli in hypnotic and hysterical anesthesia. *Johns Hopk. Hosp. Bull.*, 1930, **46**, 331–339.

49. Malmo, R. B., *et al.* Standardized pain stimulation as controlled stress in physiological studies of psychoneurosis. *Science*, 1948, **108**, 509–511.

50. McCleary, R. A., & Lazarus, R. S. Autonomic discrimination without awareness: An interim report. *J. Personality.* (In press.)

51. McGinnies, E. Emotionality and perceptual defense. *Psychol. Rev.* (In press.)

52. McDowall, R. J. The physiology of the psychogalvanic reflex. *Quart. J. exp. Physiol.*, 1933, **23**, 277–285.

53. Markbreiter, R. The effect of atropine on the emotive response. *Proc. roy. Soc. Lond.*, Series B, 1919–1920, **91**, 40–43.

54. O'Leary, W. D. The autonomic nervous system as a factor in the psychogalvanic reflex. *J. exp. Psychol.*, 1932, **15**, 767–772.

55. Peterson, F., & Jung, C. G. Psychophysical investigations with the galvanometer and pneumograph in normal and insane individuals. *Brain*, 1907, **30**, 153–218.

56. Redlich, F. C. Organic and hysterical anesthesia. *Amer. J. Psychiat.*, 1945, **102**, 318–324.

57. Richter, C. P. The sweat glands studied by the electrical resistance method. *Amer. J. Physiol.*, 1924, **68**, 147.

58. Richter, C. P. The significance of changes in the electrical resistance of the body during sleep. *Proc. nat. Acad. Sci., Wash.*, 1926, **12**, 214–222.

59. Richter, C. P. Pathological sleep and similar conditions studied by electrical skin resistance method. *Arch. Neurol. Psychiat.*, 1929, **21**, 363–375.

60. Richter, C. P. Physiological factors involved in the electrical resistance of the skin. *Amer. J. Physiol.*, 1929, **88**, 596–615.

61. Richter, C. P. Galvanic skin reflex from animals with complete transection of the spinal cord. *Amer. J. Physiol.*, 1930, **93**, 468–472.

62. Richter, C. P. Sleep produced by hypnotics studied by the electrical skin resistance method. *J. Pharmacol. exp. Therap.*, 1931, **42**, 471–486.

63. Richter, C. P., & Hines, M. Experimental production of the grasp reflex in adult monkeys by frontal lobe lesions. *Amer. J. Physiol.*, 1932, **101**, 87–88.

64. Richter, C. P., *et al.* Facial patterns of electrical skin resistance. *Johns Hopk. Hosp. Bull.*, 1942, **70**, 442–459.

65. Richter, C. P., *et al.* Hand and foot patterns of low electrical resistance. *J. Neurophysiol.*, 1943, **6**, 417–424.

66. Richter, C. P., & Whelan, F. Sweat gland responses to sympathetic stimulation by galvanic skin reflex method. *J. Neurophysiol.*, 1943, **6**, 191–194.

67. Schwartz, H. G. Reflex activity within the sympathetic nervous system. *Amer. J. Physiol.*, 1934, **109**, 593–603.

68. Schwartz, H. G. Effect of experimental lesions of the cortex on the psychogalvanic reflex in the cat. *Arch. Neurol. Psychiat.*, 1937, **38**, 308–320.

69. Seward, J. P., & Seward, G. H. The relation of galvanic skin reactions to preceding resistance. *J. exp. Psychol.*, 1935, **18**, 64–79.

70. Sidis, B. The nature and cause of the galvanic phenomenon. *J. abnorm. Psychol.*, 1910, **5**, 69–74.

71. Sidis, B., & Nelson, L. The nature and causation of the galvanic phenomena. *Psychol. Rev.*, 1910, **17**, 98–146.

72. Spiegal, E. A., & Hunsicker, W. C. Conduction of nervous impulses of

autonomic system. *J. nerv. ment. Dis.*, 1936, **83**, 252–274.

73. Thouless, R. H. The technique of experimentation on the psychogalvanic reflex phenomenon and the phenomenon of Tarchanoff. I. *Brit. J. Psychol.*, 1930, **20**, 219–240.
74. Tower, S., & Richter, C. P. Injury and repair within the sympathetic nervous system. *Arch. Neurol. Psychiat.*, 1932. **28**, 1139–1152.
75. Wang, G. H., & Richter, C. P. Action currents from the pad of the cat's foot produced by stimulation of the tuber cinereum. *Chin. J. Physiol.*, 1928, **2**, 279–283.
76. Wang, G. H., & Lu, T. W. The rate of conduction in the post-ganglionic sympathetic nervous fibers to the sweat glands in the cat's food pad. *Chin. J. Physiol.*, 1929, **3**, 335–339.
77. Wang, G. H., & Lu, T. W. Galvanic skin reflex induced in the cat by stimulation of the motor area ofthe cerebral cortex. *Chin. J. Physiol.*, 1930, **4**, 303–324.
78. Wang, G. H., & Lu, T. W. On the intensity of the galvanic skin reflex induced by stimulation of postganglionic sympathetic nerve fibers with single induction shocks. *Chin. J. Physiol.*, 1930, **4**, 393–399.
79. Wang, G. H., & Mok, W. H. The effect of hemisection of the cervical spinal cord on the galvanic skin reflex induced by cortical stimulation. *Chin. J. Physiol.*, 1931, **5**, 141–148.
80. Waller, A. D. The galvanometric measurement of "emotive" physiological changes. *Proc. roy. Soc. Lond.*, Series B, 1918, **90**, 214–217.
81. Wells, F. L., & Forbes, A. On certain electrical processes in the human body and their relation to emotional reactions. *Arch. Psychol., N. Y.*, 1911, **16**, 1–39.
82. Wells, F. L. The effect of the circulation on the electrical resistance of the skin of man. *Quart. J. exp. Physiol.*, 1927, **18**, 33–44.
83. Wenger, M. A., & Irwin, O. C. Fluctuations in the skin resistance of infants and adults and their relation to muscular processes. *Univ. Ia. Stud. Child Welf.*, 1936, **12**, 143–179.
84. White, M. M. Relation of bodily tension to electrical resistance. *J. exp. Psychol.*, 1930, **13**, 267–277.

Received August 31, 1949.

HOMOGENEITY OF VARIANCE AND THE LATIN SQUARE DESIGN[1]

ALLEN L. EDWARDS
The University of Washington

I. INTRODUCTION

The basic principles of the Latin square design are discussed in Fisher (**4**), Kendall (**9**), Goulden (**6**), Snedecor (**10**), and Tippett (**11**). Of major interest to psychologists, however, is the excellent discussion of the applications of this design to psychological research published by Grant (**6**).

In this paper we shall be concerned with the assumption of homogeneity of error variance in two particular cases treated by Grant: (1) the replication of an experiment with a series of independently drawn Latin squares; (2) the replication of an experiment by repeating the same Latin square. The first of these two cases is illustrated by the data of Bliss and Rose (**2**) presented in Table I and the second by the data of Horton (**7**) presented in Table VI.

In the typical analysis of variance problem, including the two cases mentioned above, the total sum of squares, based upon degrees of freedom equal to one less than the total number of observations, is analyzed into a number of independent sums of squares, each based upon a specified number of degrees of freedom.[2] These sums of squares and degrees of freedom are additive. Each sum of squares, when divided by its degrees of freedom, yields a mean square.

Depending upon the design of the experiment, some of the mean squares are taken as estimates of variation to be expected in random sampling from populations with a common variance (experimental error) *plus* components which reflect differences, if any are present, in population means (experimental treatments or conditions). Other mean squares are assumed to estimate only variation associated with random sampling from populations with a common variance. Thus the sums of squares and degrees of freedom for the latter are usually pooled to arrive at a single estimate of the common variance. Whenever this is done the assumption of homogeneity of mean squares to be pooled is involved and the pooling is proper only if this assumption is tenable. A test of the hypothesis of homogeneity of variance is provided by Bartlett (**1**).[3] With this brief discussion in mind, let us examine the data of Bliss and Rose, Table I.

---

[1] I am indebted to Dr. David Grant for a critical reading of this paper.

[2] Methods of calculating these sums of squares are described by Grant (**6**) and the other references cited and will not be repeated here.

[3] Other tests of the hypothesis are available also, in addition to that proposed by Bartlett. These are discussed in Johnson (**8**).

TABLE I

MG. PER CENT CALCIUM SECRETION OF DOGS GIVEN FOUR DIFFERENT PREPARATIONS OF PARATHYROID EXTRACT

LATIN SQUARES				DOGS	DAYS				
					3/15	3/25	4/5	4/15	Sum
S_1	S_2	U_2	U_1	1	13.8	17.0	16.0	16.0	62.8
U_2	U_1	S_1	S_2	2	15.8	14.3	14.8	15.4	60.3
S_2	S_1	U_1	U_2	3	15.0	14.5	14.0	15.0	58.5
U_1	U_2	S_2	S_1	4	14.7	15.4	14.8	14.0	58.9
				Sum	59.3	61.2	59.6	60.4	240.5
U_2	U_1	S_1	S_2	5	17.0	16.5	15.0	15.4	63.9
U_1	U_2	S_2	S_1	6	15.1	15.0	15.8	13.4	59.3
S_2	S_1	U_1	U_2	7	15.0	14.0	14.6	15.6	59.2
S_1	S_2	U_2	U_1	8	12.0	13.8	14.0	13.8	53.6
				Sum	59.1	59.3	59.4	58.2	236.0
S_2	U_2	S_1	U_1	9	14.6	15.4	14.0	14.8	58.8
U_1	S_1	U_2	S_2	10	13.6	15.3	17.2	15.3	61.4
U_2	S_2	U_1	S_1	11	14.4	13.8	14.4	15.0	57.6
S_1	U_1	S_2	U_2	12	15.8	15.0	15.2	15.8	61.8
				Sum	58.4	59.5	60.8	60.9	239.6
U_1	U_2	S_1	S_2	13	14.0	13.8	14.0	14.0	55.8
U_2	U_1	S_2	S_1	14	16.2	14.0	13.0	13.0	56.2
S_2	S_1	U_2	U_1	15	13.0	14.0	14.0	13.0	54.0
S_1	S_2	U_1	U_2	16	13.2	16.0	14.9	16.4	60.5
				Sum	56.4	57.8	55.9	56.4	226.5
S_1	U_1	S_2	U_2	17	14.2	14.1	15.0	14.4	57.7
U_1	S_1	U_2	S_2	18	13.0	13.4	13.8	14.0	54.2
U_2	S_2	U_1	S_1	19	15.8	16.0	15.0	15.4	62.2
S_2	U_2	S_1	U_1	20	15.2	16.2	15.0	15.3	61.7
				Sum	58.2	59.7	58.8	59.1	235.8

II. REPLICATION OF AN EXPERIMENT WITH A SERIES OF INDEPENDENTLY DRAWN LATIN SQUARES

Bliss and Rose were concerned with a problem in biological assay in which two doses, U_1 and U_2, of an unknown preparation of parathyroid extract were administered along with two doses, S_1 and S_2, of a standard preparation. The four drugs were administered on four different days, separated by the time intervals shown in the table. Measurements were made in terms of the mg. per cent calcium secretion of the dogs in response to the drugs. Using four dogs as subjects, a single 4×4 Latin square may be formed with the columns corresponding to days, the rows to dogs, and the cell entries to drugs. In conformity with the Latin square design, each drug appears once and only once in each column and each row of the 4×4 square. The experiment was replicated with a series of five independently drawn Latin squares.

The analysis of variance for the combined data of Table I, as made by Bliss and Rose, is shown in Table II. The total sum of squares is equal to 86.59 and is based upon 79 degrees of freedom. To show the components of the error sum of squares, 27.31, based upon 54 degrees of freedom, let us consider the analysis of variance of the five independently drawn Latin squares separately. For example, the analysis of variance of the first Latin square is summarized in Table III. An analysis exactly like this may be made for each of the other Latin squares. Then pooling the corresponding sums of squares and degrees of freedom,

TABLE II

ANALYSIS OF VARIANCE OF THE DATA OF TABLE I

SOURCE OF VARIATION	SUM OF SQUARES	*df*	MEAN SQUARE	*F*
Days (columns)	.94	3	.313	
Dogs (rows)	43.19	19	2.273	4.49
Drugs	15.15	3	5.050	9.98
Error	27.31	54	.506	
Total	86.59	79		

TABLE III

BREAKDOWN OF THE SUM OF SQUARES AND DEGREES OF FREEDOM FOR THE FIRST LATIN SQUARE OF TABLE I

SOURCE OF VARIATION	SUM OF SQUARES	*df*
Days (columns)	.546	3
Dogs (rows)	2.832	3
Drugs	4.756	3
Error	3.160	6
Total	11.294	15

we would have a sum of squares for days based upon 15 degrees of freedom, a sum of squares for dogs based upon 15 degrees of freedom, a sum of squares for drugs based upon 15 degrees of freedom, and a sum of squares for error based upon 30 degrees of freedom, making a total of 75 degrees of freedom. It will be found that the sum of these sums of squares will not equal the total sum of squares for the combined data, 86.59. The difference is equal to 7.69 and this sum of squares is based upon the remaining 4 degrees of freedom and is associated with differences among the five Latin squares. Thus, we have the analysis of the total sum of the squares and degrees of freedom as shown in Table IV.

TABLE IV

POOLED SUMS OF SQUARES AND DEGREES OF FREEDOM FOR THE FIVE LATIN SQUARES OF TABLE I

SOURCE OF VARIATION	SUM OF SQUARES	*df*
Days (columns)	2.62	15
Dogs (rows)	35.50	15
Drugs	22.57	15
Error	18.21	30
Latin squares	7.69	4
Total	86.59	79

Note that the error sum of squares, 18.21, based upon 30 degrees of freedom, is itself composed of five components, each based upon 6 degrees of freedom, and that each of these yields a mean square which is an estimate of experimental error. Thus, the assumption of homogeneity of the error variance is involved, if they are to be pooled. Direct calculation of these mean squares shows that they are equal to .527, .332, .700, 1.374, and .102, respectively, for the five Latin squares. Applying Bartlett's test, we find that χ^2, as given in line 4 of Table V, is equal to 9.41 with 4 degrees of freedom.[4] This value is of borderline significance, but since the χ^2 obtained in line 4 of Table V tends to be slightly biased in that it overestimates the "true" value, we may obtain a correction factor, suggested by Bartlett, and shown in line 5. Then the "corrected" value of χ^2, is given in line 6, as equal to 8.79. For 4 degrees of freedom, $P_{.10}$ is 7.779 and we may consider the null hypothesis tenable.[5] There is

---

[4] The degrees of freedom available for evaluating χ^2 in the test of homogeneity of variance will be one less than the number of mean squares tested.

[5] What would a significant value of χ^2 mean in this instance? It would indicate that the experimental technique is not very reliable, i.e., that the error mean squares from the

TABLE V

TEST OF HOMOGENEITY OF VARIANCE OF THE ERROR MEAN SQUARES FROM THE FIVE LATIN SQUARES OF TABLE I

LATIN SQUARES	SUM OF SQUARES	*df*	MEAN SQUARE	LOG MEAN SQUARE
I	3.160	6	.527	$\bar{1}.72181$
II	1.990	6	.332	$\bar{1}.52114$
III	4.200	6	.700	$\bar{1}.84510$
IV	8.246	6	1.374	.13799
V	.614	6	.102	$\bar{1}.00860$
Sum	18.210	30	3.035	$\bar{2}.23464$

(1) $\frac{\Sigma s^2}{5} + \frac{3.035}{5} = .6070$; $\log .6070 = \bar{1}.78319$

(2) $(5)(\bar{1}.78319) = \bar{2}.91595$

(3) $\text{Diff.} = \bar{2}.91595 - \bar{2}.23464 = .68131$

(4) $\chi^2 = (2.3026)(6)(.68131) = 9.41$; $df = 4$; $P_{.05} = 9.488$

(5) $C = 1 + \frac{6}{(3)(5)(6)} = 1.07$

(6) $\chi^2_{\text{corrected}} = 9.41/1.07 = 8.79$; $df = 4$; $P_{.10} = 7.779$

no reason why we should not pool the error sums of squares from the separate Latin squares along with their degrees of freedom to arrive at a common estimate for the error variance based upon 30 degrees of freedom.

We have not as yet arrived at the error mean square of .506 based upon 54 degrees of freedom as shown in Table II. It can be demonstrated, however, that the additional components are the two interactions: (1) Latin squares × days and (2) Latin squares × drugs. For example, the sum of squares for days, 2.62, based upon 15 degrees of freedom, can be analyzed into two components: (1) a sum of squares for days, .94, with 3 degrees of freedom and (2) a sum of squares for the interaction between Latin squares and days, 1.68, based upon 12 degrees of freedom. Similarly, the sum of squares for drugs, 22.57, based upon 15 degrees of freedom, can be analyzed into two components: (1) a sum of squares for drugs, 15.15, with 3 degrees of freedom and (2)

---

various replications of the experiment vary more than can reasonably be attributed to chance or random sampling from a common population. If we reject the hypothesis of random sampling from a population or populations with a common variance, then we obviously have no basis for combining the separate values.

a sum of squares for the interaction between Latin squares and drugs, 7.42, based upon 12 degrees of freedom.

Now the five Latin squares, if randomly drawn, constitute a random sample from the population of 4×4 Latin squares. With a reliable experimental technique there is no reason for believing that the interactions involving the Latin squares are anything more than estimates of experimental error also. Whether these two assumptions are tenable or not may be determined by testing the two interaction mean squares against the known estimate of error obtained from the Latin squares. Thus, for the Latin squares×drugs interaction, $F=(7.42/12)/(18.21/30)=1.02$. Since an F of 2.84 is required for significance at the 2 per cent level and an F of 2.09 for significance at the 10 per cent level for 12 and 30 degrees of freedom, the obtained value of 1.02 is not significant and the null hypothesis may be regarded as tenable.[6]

The mean square for the Latin squares×days interaction is $1.68/12=.140$ and this is smaller than the error mean square $18.21/30=.607$. The test of significance may be made by taking $F=.607/.140=4.34$ with 30 degrees of freedom for the numerator and 12 for the denominator. The tabled values of F at the 2 per cent and 10 per cent levels of significance for 30 and 12 degrees of freedom are 3.70 and 2.46, respectively. Thus, the Latin squares×days mean square must be regarded as significantly smaller than the error mean square as obtained from the Latin squares. Since this interaction would ordinarily not be expected to be smaller than the error mean square, except by chance, this particular result may be regarded as one of those unusual cases that do sometimes happen by chance.[7]

We thus see that the error sum of squares as given by Bliss and Rose, Table II, is obtained by adding the pooled error sums of squares from the five Latin squares with the two interaction sums of squares involving the Latin squares to get $18.21+7.42+1.68=27.31$. Pooling the degrees of freedom associated with these sums of squares gives $30+12+12=54$. The error mean square of their analysis is thus given by $27.31/54=.506$. Strictly speaking, the best *common* estimate of error which can be obtained from the data is that based upon the pooled error sums of squares from the separate Latin squares plus the sum of squares for the inter-

---

[6] Since this is a two-tailed test of significance, the probabilities associated with the tabled values of F as given by Snedecor (**10**) must be doubled.

[7] In the absence of any real interaction, the average value of the Latin squares×days interaction may be expected to be that of the error mean square. In which case in any particular experiment it may, of course, be smaller by chance. An interaction between the Latin square and the day variable, if it existed, would serve to increase the average value of this mean square over that of the error mean square.

interaction which is equal to 376.19 and is based upon 16 degrees of freedom and (2) a sum of squares for the *pooled* interactions of subjects ×trials for the separate orders of presentation which is equal to 576.40 and is based upon 80 degrees of freedom. This latter sum of squares can, in turn, be analyzed into five components: the subjects×trials interactions within each of the five sequences, each based upon 16 degrees of freedom.

The trials×sequences interaction sum of squares can also be further analyzed into two components: (1) a sum of squares for the screen size which is equal to 272.61 and is based upon 4 degrees of freedom and (2) a sum of squares for the residual error of the Latin square design which is equal to 103.58 and is based upon 12 degrees of freedom. The successive breakdowns in this analysis are shown in Table VII.

TABLE VII

SUCCESSIVE BREAKDOWNS OF THE TOTAL SUM OF SQUARES AND DEGREES OF FREEDOM FOR THE DATA OF TABLE VI

SOURCE OF VARIATION	SUM OF SQUARES			*df*	MEAN SQUARE
(1) Trials	42.21	42.21	42.21	4	10.55
(2) Total between subjects	2,475.49				
(2a) Order of presentation		60.69	60.69	4	15.17
(2b) Between S's of same order		2,414.80	2,414.80	20	120.74
(3) Total subjects×trials	952.59				
(3a) Order×trials		376.19			
(3aa) Screen size			272.61	4	68.15
(3ab) Latin square error			103.58	12	8.63
(3b) Pooled S's same order×trials		576.40			
(3ba) S's order I×trials			116.40	16	7.275
(3bb) S's order II×trials			171.76	16	10.735
(3bc) S's order III×trials			70.56	16	4.410
(3bd) S's order IV×trials			98.64	16	6.165
(3be) S's order V×trials			119.04	16	7.440
(4) Total for experiment	3,470.29	3,470.29	3,470.29	124	

Let us test first for the homogeneity of variance of the five mean squares involving interactions between subjects with a given order of presentation and trials. Bartlett's test is applied to these values in Table VIII and we find that χ^2 is equal to 3.35 with 4 degrees of freedom.[8] This is not a significant value since P is slightly larger than .50.

---

[8] There is no need to apply the correction suggested by Bartlett to this value, since the correction will only tend to reduce the size of the obtained value.

TABLE VIII

TEST OF HOMOGENEITY OF VARIANCE OF THE INTERACTION MEAN SQUARES FOR SUBJECTS GIVEN THE SAME ORDER OF PRESENTATION AND TRIALS

SOURCE OF VARIATION	SUM OF SQUARES	*df*	MEAN SQUARE	LOG MEAN SQUARE
Subjects order I × trials	116.40	16	7.28	.86213
Subjects order II × trials	171.76	16	10.74	1.03100
Subjects order III × trials	70.56	16	4.41	.64444
Subjects oraer IV × trials	98.64	16	6.16	.78958
Subjects order V × trials	119.04	16	7.44	.87157
Sum	576.40	80	36.03	4.19872

(1) $\frac{\Sigma s^2}{5}=\frac{36.03}{5}=7.21$; $\log 7.21=.85794$

(2) $(5)(.85794)=4.28970$

(3) Diff. $=4.28970-4.19872=.09098$

(4) $\chi^2=(2.3026)(16)(.09098)=3.35$; $df=4$; $P_{.05}=9.488$

Since these mean squares may be judged homogeneous, the sums of squares may be pooled along with the associated degrees of freedom to arrive at a mean square of 7.20 based upon 80 degrees of freedom.

Now the mean square 7.20, based upon 80 degrees of freedom, under the conditions of the experiment and in the absence of significant interactions between subjects with a given order of presentation and trials, should estimate the same quantity as the error mean square obtained from the Latin square design, 8.63. A test for the homogeneity of these two values is made by taking $F=8.63/7.20=1.20$ with 12 and 80 degrees of freedom. By reference to the table of F we find that this is not a significant value. Hence these two sums of squares and associated degrees of freedom may be pooled to arrive at the common estimate, 7.39, based upon 92 degrees of freedom. We thus arrive at the analysis shown in Table IX.

Note that two error terms are indicated: one based upon the variation among independent randomly assigned subjects (inter-subject variation), tested under the same order of presentation, and the second based upon the variation within the same subjects (intra-subject variation) or, in other words, upon repeated observations of the same subjects.[9] Thus, the mean square for orders of presentation may be tested against the error mean square 120.74, and the mean squares for size of

---

[9] The analysis of variance applied to other experimental designs involving repeated observations of the same subjects is described by Edwards (3).

TABLE IX

ANALYSIS OF VARIANCE OF THE DATA OF TABLE VI

SOURCE OF VARIATION	SUM OF SQUARES	*df*	MEAN SQUARE	*F*
Independent observations				
Order of presentation	60.69	4	15.17	—
Between S's of same order	2,414.80	20	120.74	
Total between subjects	2,475.49	24		
Correlated observations				
Screen size	272.61	4	68.15	9.22
Trials	42.21	4	10.55	1.43
Pooled error	679.98	92	7.39	
Total within subjects	994.80	100		
Total for the experiment	3,470.29	124		

screen and for trials may be tested against the error mean square 7.39. Making these tests, we find that the screen size is the only significant variable in the experiment.

IV. SUMMARY

In replicating an experiment with a series of independently drawn Latin squares, the pooling of the residual error sums of squares from the separate Latin squares involves the assumption of homogeneity of variance. This assumption may be tested by means of Bartlett's test. If the hypothesis of homogeneity of variance is not tenable, this is an indication that the experimental technique is not reliable; that the error variances from the separate replications differ more than can be attributed to random sampling from a common population. A contributing factor may be the presence of interactions in some of the Latin squares between the row and column variables.[10] Since the column variable will ordinarily correspond to successive trials or periods of testing and, under any circumstances, will be the same from Latin square to Latin square, an examination of the row variable (which will ordinarily correspond to different orders of presenting the experimental treatments as well as to subjects) may throw some light upon the nature of the interaction. It should be emphasized, however, that replication of an experiment with a series of independently drawn Latin squares permits no separa-

[10] See the discussion by Grant (6).

tion of the variation attributable to the various orders of presentation of the experimental treatments and differences among subjects presented with the same order.

By replicating the same Latin square, on the other hand, it is possible to separate the row variation into two components. One part will correspond to the variation attributable to the particular orders of presentation of the experimental treatments and the other will measure the variation attributable to randomly assigned subjects presented with the same order. The mean square for the latter may be used to test the significance of the particular orders of presenting the experimental treatments. This design also permits the calculation of the separate interactions for subjects × trials for each order of presentation. These interaction mean squares may be tested for homogeneity and the sums of squares and associated degrees of freedom may be pooled if the null hypothesis is tenable.

In the absence of significant interactions among subjects tested under the same order and trials, the pooled mean square based upon these interactions should estimate the same quantity as that derived from the error sum of squares of the Latin square design. A test of homogeneity may be made for these two mean squares and, if the null hypothesis is tenable, the two may be pooled to arrive at a common estimate of error. This error term may be used to test the significance of the experimental treatments or cell entries of the Latin square design and also to test the significance of the column variable (days or trials).

BIBLIOGRAPHY

1. BARTLETT, M. S. Some examples of statistical methods of research in agriculture and applied biology. *J. R. statist. Soc. Suppl.*, 1937, **4**, 137–170.
2. BLISS, C. I., & ROSE, C. L. The assay of parathyroid extract from the serum calcium of dogs. *Amer. J. Hyg.*, 1940, **31**, 79–98.
3. EDWARDS, A. L. *Experimental design in psychological research.* New York: Rinehart, 1950.
4. FISHER, R. A. The design of experiments. (3rd ed.) Edinburgh: Oliver & Boyd, 1942.
5. GOULDEN, C. H. *Methods of statistical analysis.* New York: Wiley, 1939.
6. GRANT, D. A. The Latin square principle in the design and analysis of psychological experiments. *Psychol. Bull.*, 1948, **45**, 427–442.
7. HORTON, G. P. Unpublished data. Seattle: University of Washington, 1949.
8. JOHNSON, P. O. *Statistical methods in research.* New York: Prentice-Hall, 1949.
9. KENDALL, M. G. *The advanced theory of statistics.* Vol. II. (2nd ed.) London: Griffin, 1948.
10. SNEDECOR, G. W. *Statistical methods.* (4th ed.) Ames, Iowa: State College Press, 1946.
11. TIPPETT, L. H. C. *The methods of statistics.* (3rd ed.) London: Williams & Norgate, 1941.

Received January 8, 1950.

A QUALIFICATION IN THE USE OF ANALYSIS OF VARIANCE

WILSE B. WEBB AND VERNON W. LEMMON
Washington University

In problems involving more than two conditions (groups, variables, etc.) the increasing use of the analysis of variance in the consideration of such data has been apparent. That this statistical technique has become a typical method of analysis in the problems of psychology and education is evidenced by its inclusion in most of the recent statistics texts in this field (**1**, **3**, **4**, **5**, **6**). For a particular and common experimental design in this area, however, the authors have noted that the technique presents certain limitations. In fact, as will be noted, the use of this method of analysis may lead to fallacious conclusions. Most importantly, the texts, which must serve as guides to the experimentalist, not only have failed to note this limitation but, in nearly every instance, foster this error.

A typical experimental design in psychological research involves the effect of levels within an ordered continuum of events on some behavior. Examples are numerous: levels of drive on learning or response; some magnitude of stimulus intensity as it affects learning or response; size of print and reading proficiency; age of subject and various behavior variables, etc. These typical problems involve a general question: does the ordered (independent) variable have some effect on the response (dependent) variable. In executing such a problem, random groups are selected and the various levels of the ordered variable are introduced and the resultant changes are measured. Under such circumstances the analysis of variance has been increasingly used to test for the existence of a significant effect. In the simple case the analysis is in the form of the determination of a methods variance and a within group variance. If the F (the ratio of the methods to the within groups variance) resulting from these two variances is significant the experimenter draws conclusions concerning the operation of a real effect by the independent variable. If, however, the F is insignificant the experimenter concludes that no relationship exists and no further analysis is performed.

It is not surprising that the procedure cited above should be the typical one. The experimenter, keen on his experiment but rusty on his statistics, will seek the nearest statistical text available. In doing so he will find the following statements.

After and only after it has been found that the overall F *is significant, can one* *safely use the* t *technique to test the significance of the difference between any two of the group means* (6).[1]

If F *is not significant there is no point in going further as no mean difference can be significant.* But if *F* is significant we may proceed to calculate C. R.'s for the differences between the column means (3).

The answer is that the F-test tells us whether it is worthwhile to test the individual differences at all. *If the* F*-test had not been proven significant, we would have known at once that all observed differences in methods means could be due to chance alone. In that case it would not only have been unnecessary but decidedly improper to apply the* t*-test to individual differences.* [A footnote is appended as follows:] This is particularly true if all of the methods are qualitatively similar, and if the experiment was not designed for any special comparison of those of the methods. An exception to this rule may be defended, however, if one of the methods exhibited marked qualitative differences from all of the others. [Here an example is given of four similar procedures in contrast to a dissimilar procedure. The *F* for the five procedures was short of significance but the *t*-test was used to compare the obviously different procedure with the group mean of the other four procedures.] This procedure was legitimate in this case because the special comparison was suggested by the qualitative characteristics of the methods and not just by the inspection of the final means. Had all five methods been qualitatively similar, however, it would not have been legitimate to select the method with the highest (or lowest) criterion mean for special comparison with one or with a combination of the others. *Design of Experiments,* page 65 [2] (5).

What if the value of *F* which is obtained fails to be significant? May we then use the *t*-test to test the difference between the pairs of means? [An example is then given of choosing the highest and lowest mean in a group of ten and receiving a significant *t* in the face of an insignificant *F* and the question is posed as whether the *F* test and the *t* test are inconsistent.] When we examine the hypothesis tested by the *t* test *we see the fallacy involved in this comparison.* The hypothesis tested by *t* is that the two sample means have been drawn at *random* from the same population. We have selected the largest and the smallest for comparison (1).

This author then cites Fisher's warning that comparisons suggested *after* data are in are open to suspicion, and proposes under these circumstances that a higher level of probability be required. He concludes with a quotation from Fisher that such *t*'s are to be used only as suggestions for further experimentation, in which they can be deliberately tested.

In Fisher's text, referred to by Lindquist and by Edwards, we find,

If the yields of the different varieties in the experiment fail to satisfy the test of significance they will not often need to be considered further, for the results, as so far tested, are compatible with the view that all differences observed in the experiment are due to variations in the fertility of the experimental area, and this is the simplest interpretation to put upon the results. If, however, a significant value of

---

[1] Italics in the following quotes are those of the present writers.

> *z* has been obtained the null hypothesis has been falsified, and may therefore be set aside.
>
> When the *z* test does not demonstrate significant differentiation, much caution should be used before claiming significance for special comparisons. Comparisons, which the experiment was designed to make, may, of course, be made without hesitation. It is comparisons suggested subsequently, by a scrutiny of the results themselves, that are open to suspicion; for if the variants are numerous, a comparison of the highest with the lowest observed value, picked out from the results, will often appear to be significant, even from undifferentiated material. Properly, such unforeseen effects should be regarded only as suggestions for future experimentation, in which they can be deliberately tested (2).

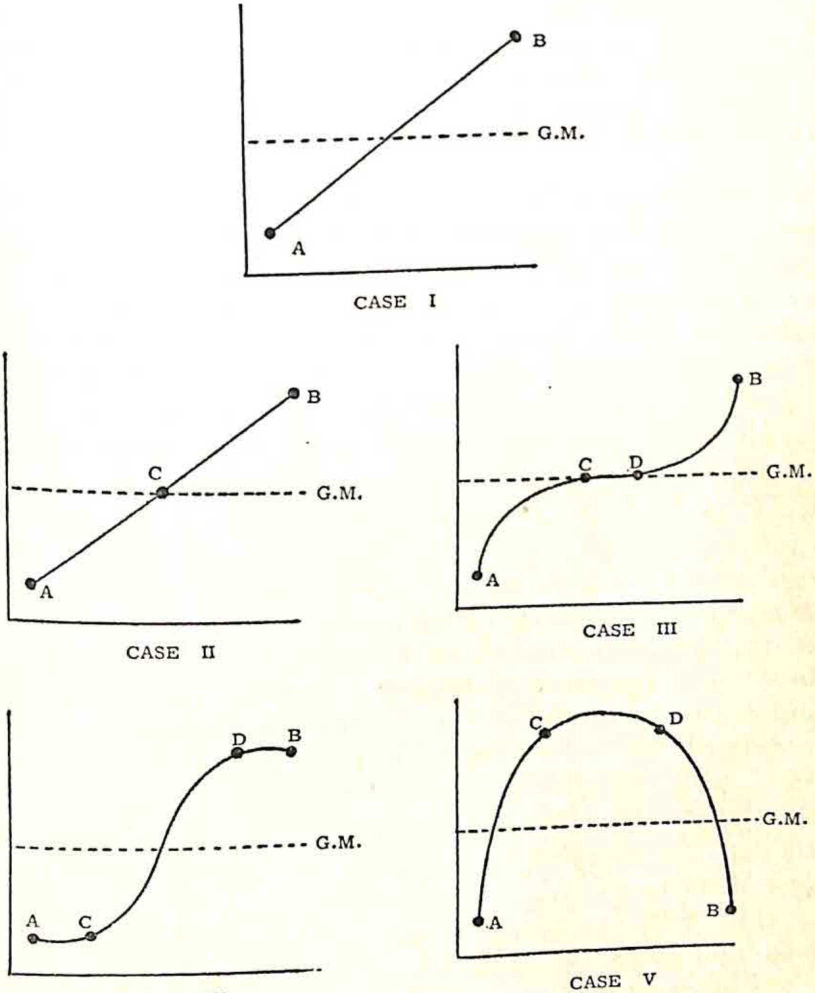
Finally, the reader will find no reference to the problem one way or the other in Guilford's discussion of analysis of variance (4).

In general then, although Lindquist and Edwards (quoting Fisher) do note certain exceptions, the unwary, or at least, statistically rusty, experimentalists would most likely conclude from an insignificant *F* ratio that further analysis would be inappropriate. This is specifically indicated in two of the general statistical texts and emphasized in the others. Lindquist and Edwards, in their noted exceptions, quote Fisher in a manner which seems more to support the non-use of *t* than its uses. Only in Fisher's *Design of Experiments* (reputedly difficult for the psychologist and not frequently available for his use) is found, in the opinion of the authors, an appropriate evaluation of such circumstances. To quote Fisher, "Comparisons, which the experiment was designed to make, may, of course, be made without hesitation " (2). Even here the unequivocal statement is bound, under circumstances, by cautions and we are not certain that the circumstances cited below were considered. It is the contention of the authors that, in the cases cited below, which constitute a considerable proportion of psychological research, the *F*-test and the *t*-test do lead to incompatible results as a function of the experimental design itself, and most importantly, regardless of the caution of the experimenter he will find circumstances where "true" differences found by one method will result in "chance" difference indications by the other. Where such circumstances exist it may be suggested that other methods are more appropriate.

Several hypothetical cases will now be investigated. The notation used will be that given by Lindquist (5): r = the number of groups and n = the number of cases per group.

Case I. Assume two groups, A and B, with 11 cases each and equal variances, which are related to some other variable X by a straight line function. We will make the further assumption that the means of A and B are just significantly different at the 5% level as shown by the *t* test.

For $r(n-1)=2\times10=20$ degrees of freedom, this would imply a t value of 2.086. Since there are only two groups, the F value for 5% significance should equal t^2, or 4.35. This may be checked from the F table, using $r-1=1$ *df* for the numerator (or between groups variance), and $r(n-1)=20$ *df* for the denominator (or within groups variance). This is the basic case, for which the t and F values are just significant. It is illustrated schematically in Fig. 1, the dotted horizontal line representing the grand mean of the two groups.



CASE I

CASE II

CASE III

CASE IV

CASE V

FIG. 1. ILLUSTRATIONS IN WHICH THE F-TEST AND THE t-TEST LEAD TO CONFLICTING ESTIMATES OF THE SIGNIFICANCE OF RELATIONSHIPS AMONG VARIABLES

Case II. Groups A and B are retained, but a third group, C, is used to investigate the value of the function for a value of the X variable half-way between the values corresponding to A and B. This group is assumed to have the same number of cases, and the same variance, as A or B. Since the function is assumed to be rectilinear the mean of C would lie on the grand mean of all three groups.

The *t* value of the difference between the means of A and B would not be changed in any way by the addition of group C. However, if we should investigate the differences between all three means by the *F* test, we would get a value differing from that of Case I. The within groups sum of squares would be increased by 50%, as compared with Case I, but the *df* for this variance would also be increased by 50%, so the denominator of the *F* ratio would not be changed. The between groups sum of squares would not be contributed to by group C, since its mean coincides with the grand mean, but the *df* would be increased from 1 to 2. Thus the numerator of the *F* ratio, and consequently the ratio itself, would be half the value for Case I, i.e. it would now be 2.175. The *F* ratio required for significance is given by the table (for *df* 2 and 3) as 3.32. This case shows a non-significant *F* for all three groups together, but a significant *t* for the difference between A and B. The authorities previously cited indicate that the *t* test ought not be tried when the *F* test is not significant, since under such circumstances a significant *t* might be due to an unusually large chance difference. However, in Case II points A and B differ mainly because they are on the straight line function; their difference is no more due to chance than it was in Case I, in which the use of the *t* test would not have been questioned. It is not likely that the investigation of an additional point in the middle of the line would lessen the significance of the difference between the two end points.

Case III. Groups A and B are taken as before, but are assumed to be related to variable *X* by a sort of reversed "S" curve with steep ends and a nearly horizontal center portion, as shown in Fig. 1. Similar groups C and D are taken on this center portion, and for the sake of simplicity their means are assumed to coincide with the grand mean. The value of *df* would now be 40. The between groups variance would be the same as for Case I, although the changed, but the *df* by which it is divided would now be 3. The *F* ratio would therefore be 1.45, or one third the value for Case I. The *F* required for significance with *df* 3 and 40 is 2.84. Thus the *F* is even further from significance than in the previous case, while the *t* value for group means A and B remains unchanged. The same arguments used for Case II apply here.

Case IV. The same four groups of Case III are used, but the functional relationship to *X* is assumed to be represented by an "S" curve with nearly horizontal ends and a steep center portion. Groups C and D are taken very near the ends of the curve, and assumed to have means equal to those of A and B respectively (see Fig. 1). The within groups variance is unchanged, with $df=40$. Groups C and D would in this case contribute to the between groups sum of squares, which would be doubled, while the df would be 3. Thus the F ratio would be 2.90, while the F required for significance would be 2.84, as in Case III. Here both F, and the t value for A and B, would be significant. The change in F from Case III is due to the difference in the type of functional relationship to X, and to the position of the points chosen for investigation between points A and B.

Case V. This results from a rearrangement of the points of Case IV. The curve of the function is humped, with A and B being at the lower ends, and both having the mean formerly assigned to A. C and D, near the top of the curve, both have the mean formerly assigned to B. In this case the F ratio would be the same as for case IV, and therefore significant, but the t value for the difference between A and B would not be significant, since A and B now have no difference between their means. Yet even this lack of difference should not be called insignificant, since A and B are points on the curve of a definite function.

The five cases examined above show that both significant and non-significant t values can occur in conjunction with either significant or non-significant F ratios. In each case the location of the various means was not a matter of chance, but was due to a functional relationship to some other variable. Thus the various differences possessed a certain kind of significance, in spite of the erratic behavior of the t and F ratios. The latter were shown to depend largely on the shape of the curve of the function, and on the number and location of the points on the curve which were chosen for investigation.

It is suggested that in cases of functional relationship, the usual strictures on the use of the t and F tests do not apply. In fact, in such cases it is doubtful whether either of these tests yields the kind of information about significance which is needed. It would probably be more appropriate to test for the significance of regression or trend.

SUMMARY

The authors have noted in specific and common experimental designs in psychological problems that the analysis of variance may lead to results incompatible with the true resultant relations obtained. These are instances of the effect of ordered independent variables upon the dependent variable where a specific functional relation results. In such instances significant t's may be overlooked because of insignificant F ratios and insignificant t's may be considered significant on the basis of significant F's. It was further noted that frequently the resultant error was explicitly or implicitly condoned by standard statistic texts in psychology.

BIBLIOGRAPHY

1. Edwards, Allen. *Statistical analysis.* New York: Rinehart, 1946.
2. Fisher, R. A. *The design of experiments.* Edinburgh: Oliver and Boyd 1935.
3. Garrett, H. E. *Statistics in psychology and education.* New York: Longmans, Green, 1947.
4. Guilford, J. P. *Fundamental statistics in psychology and education.* New York: McGraw-Hill, 1942.
5. Lindquist, E. F. *Statistical analysis in educational research.* Cambridge, Mass: Houghton Mifflin, 1943.
6. McNemar, Quinn. *Psychological statistics.* New York: John Wiley, 1949.

Received August 8, 1949.

A SURVEY OF AUDIO-VISUAL AIDS FOR TEACHING OF CLINICAL PSYCHOLOGY[1]

ELIAS KATZ

U. S. Naval Disciplinary Barracks, San Pedro, California

The objectives of this study were (1) to determine the present use of audio-visual aids (motion pictures, recordings, slides, etc.) in teaching of clinical psychology, and (2) to elicit suggestions as to needed audio-visual aids for teaching of clinical psychology.

A questionnaire to obtain these data was sent in January and February 1949 to the 43 universities offering approved graduate training in clinical psychology, as listed in the report of the Committee of Training in Clinical Psychology of the American Psychological Association (**4**, p. 318).

A total of 50 replies was received from instructors in 30 universities. In most instances, these replies represented the actual questionnaires returned by individual instructors. In three instances, however, a representative of the psychology department wrote a letter presenting the combined responses of the individuals involved.

It must be stressed that this survey does not claim to sample adequately the entire group of instructors of clinical psychology in the universities accredited by the Committee on Training in Clinical Psychology at the time that this questionnaire was issued.

AUDIO-VISUAL AIDS USED IN TEACHING OF CLINICAL PSYCHOLOGY

Among the 50 responses, some 14 indicated that they had used no audio-visual aids (motion pictures, recordings, slides, etc.) in teaching courses in clinical psychology. Of those who had made such use, answers ranged from one or two items, to appreciable lists of motion pictures and other materials used. In one instance, the response consisted of a comprehensive listing of audio-visual aids (**2**) used as a source for the teaching of clinical psychology at that institution.

The audio-visual aids reported used were motion pictures, slides, charts and recordings. No reference was made to silent or sound film strips.

Among the lantern slides reported used were "Rorschach slides," "TAT slides," slides showing "diagrams of the mechanisms of neurosis, psychosis, socialization," "Mental Deficiency," and "Slides of my own data, MMPI's, Rorschach, Bender-Gestalt, etc."

---

[1] This study was conducted under the direction of the Committee on Teaching of Clinical Psychology, Division of Clinical and Abnormal Psychology, American Psychological Association. Grateful acknowledgment is made to the instructors of clinical psychology who submitted answers to the questionnaire. The opinions expressed herein are the writer's, and do not represent those of the Navy Department, nor of the American Psychological Association.

Among the charts used were charts showing "occupations, distribution of working populations," and a "large scattergram chart for diagnostic intelligence testing."

Recordings were reported used for Rorschach training, for speech correction work, and for training in interviewing and counseling.

The motion pictures reported used in teaching courses in clinical psychology have been listed alphabetically in the Appendix.

In the questionnaire, it was requested that the courses in which the audio-visual aid was used be indicated. In the first column of Table I, motion pictures used have been roughly grouped according to courses in the curriculum proposed by the Committee on Training in Clinical Psychology (4). This classification is admittedly gross, since it was not

TABLE I

MOTION PICTURES USED AND AUDIO-VISUAL AIDS NEEDED BY 50 INSTRUCTORS OF CLINICAL PSYCHOLOGY IN 30 UNIVERSITIES ACCREDITED BY THE COMMITTEE ON TRAINING IN CLINICAL PSYCHOLOGY (JANUARY AND FEBRUARY 1949), CLASSIFIED BY TYPES OF COURSES

The courses listed are based upon the proposed curriculum of the Committee on Training in Clinical Psychology. The number of instructors using each film is indicated in parentheses. Note that a given film may be used in different types of courses.

MOTION PICTURES USED	AUDIO-VISUAL AIDS NEEDED
General, Physiological, Comparative Psychology	
"An Introduction to Clinical Neurology" (1)	"A film showing the nature of science and the nature of psychology as science." "Functions of autonomic nervous system, and physiological effects."
Developmental Psychology, Individual Differences	
"Child Development Series" (1) "Conflict Situations in Childhood" (1) "Meeting the Emotional Needs of Childhood" (1) "This Is Robert" (1) "You and Your Child" (1) "Emotional Health" (1) "Shades of Gray" (1)	
Experimental Dynamic Psychology	
"Dynamics of an Experimental Neurosis" (2)	

(Continued on next page)

TABLE I—(*continued*)

MOTION PICTURES USED	AUDIO-VISUAL AIDS NEEDED
Psychopathology	
"Behavior in Hypnotic Regression" (1) "Convulsive and Allied Conditions" (1) "Dynamics of an Experimental Neurosis" (1) "Psychoneuroses" (1) "Electroshock Therapy in Ambulatory Practice" (1) "The Feeling of Rejection" (2) "The Feeling of Hostility" (1) "Let There Be Light" (1) "Mental Defectives" (1) "Psychiatry in Action" (4) "Symptoms in Schizophrenia" (2) "Treatment of Mental Disorders" (1)	"Depiction of various behavior disorders." (7) "Etiology of psychotic behavior." "Carried-over aggressive resentments, based on case material." "Mental mechanisms: repression and aggression." "Etiology of anti-social behavior." "Neurosis in alcoholics." "Etiology of neurotic behavior."
Observational Techniques	
"Experimentally Produced Neurotic Behavior in Rats" "Grief" (1)	"Different kinds of qualitative vs. quantitative responses to test situations." "Children of very young age (i.e., difficult to obtain 'in the flesh') in a variety of behavior." "Such as S.G. Estes' study, to aid in teaching the clinical 'pick up' of minimal cues from posture, gait, faces, expressive movements."
Survey of Clinical Psychology	
"Balloons: Aggressive and Destructive Games" (1) "Care of the Sick and Injured: The N.P. Patient" (1) "The Feeling of Rejection" "An Introduction to Clinical Neurology" (2) "Paranoid State and Deterioration Following Head Injury" (1) "Shades of Gray" (1) "Symptoms in Schizophrenia" (1) "This Is Robert" (1)	"Demonstration of clinical techniques, diagnostic and therapeutic." (1)
Methods of Case Study, Case Analysis, Interviewing	
"Psychological Implications of Behavior During the Clinical Visit" (1)	"Interview techniques." (2) "Case history and interviewing techniques."

TABLE I—(*continued*)

MOTION PICTURES USED	AUDIO-VISUAL AIDS NEEDED
Verbal "Intelligence" Tests	
"Aptitudes and Occupations" (1) "The Feebleminded" (1)	"Administrative technique of Wechsler-Bellevue." "Interpretation of Wechsler-Bellevue." "Administrative techniques of Revised Stanford-Binet." "Interpretation of Revised Stanford-Binet." "Merrill-Palmer test administration." "Infant Scales." "A pre-school performance test."
Motor Tests	
"Motor Disorders in Nervous Diseases" (1)	
Clinical Tests of Psychological Deficit, Aphasia, Conceptualization	
	"Various types of mental defectives." (2)
Projective Tests	
"Balloons: Aggressive and Destructive Games" (3) "Clinical Psychology and Hypnosis" (1) "Frustration Play Techniques" (1) 'This Is Robert" (1)	"Rorschach techniques and interpretation." "TAT administration and interpretation." "Bender-Gestalt administration and interpretation." "Organic patients on Goldstein-Scheerer, Sorting and other tests.'
Therapeutic Theory and Methods	
	"Orienting the student to the service of the clinic, and demonstrating the operation, illustrating professional principles guiding interpersonal relations."
Remedial Aspects of Special Disabilities	
	"Films on Occupational Therapy."
Techniques of Guidance and Counseling	
"The Feeling of Rejection" (1) "Narcosynthesis" (1)	"Vocational counseling." "Non-directive adjustment counseling." "Demonstration of various counseling techniques."

TABLE I—(*continued*)

MOTION PICTURES USED	AUDIO-VISUAL AIDS NEEDED
Personality Therapy	
"Dynamics of an Experimental Neurosis" (1) "The Feeling of Hostility" (1) "The Feeling of Rejection" (1) "Narcosynthesis" (1) "Psychiatry in Action" (1) "Treatment of Mental Disorders" (1)	"Shock therapies." "Standard psychoanalysis." "Better films on hypnosis." "Well handled therapeutic interviews of all kinds for use in teaching psychotherapy." "Play Therapy sessions." "Some skillfully done shots of successive therapeutic interviews for fairly simple problems by therapists using different approaches." "Neurosis: background factors and treatment steps, based on case material."
Group Therapy	
	"Group Therapy" (3) "Psychodrama" (2)
Research in Dynamic Psychology	
	"Procedures in experimental studies of learning, perception, emotion, reasoning, etc., in various classes of psychiatric and feeble-minded persons, with special attention to details of performance."
Introduction to Clinical Medicine	
"Clinical Psychology and Hypnosis" (1) "Dynamics of an Experimental Neurosis" (1) "Some Examples of Hypnotic Behavior" (1)	"Prefrontal lobotomy in the treatment of mental disorders." "Treatment of mental disorders." "Methodology of clinical neurological examinations, gross neurological signs and their significance."
Influence of Culture on Personality	
	"Typical socialization procedures used in a primitive society, on the order of the Bateson-Mead photographs of Balinese."

possible to determine whether courses of the same name taught at different universities were equivalent, or whether the Committee's proposed courses were equivalent to those being offered. Nevertheless, the findings indicated some tendencies which may be of interest.

AUDIO-VISUAL AIDS NEEDED

Among the 50 responses, 11 instructors stated that they felt no need for any audio-visual aids in their teaching. Six of these 11 were among the 14 who had not previously used audio-visual aids in their work.

The responses of those who expressed needs for audio-visual aids in their teaching of clinical psychology have been presented in the second column of Table I, again grouped according to courses in the curriculum proposed by the Committee on Training in Clinical Psychology (4).

SUMMARY

A questionnaire to determine present use of and needs for audio-visual aids was sent in January and Feburary, 1949 to 43 universities offering graduate training in clinical psychology. Fifty replies were received from 30 universities. It was not felt that this was an adequate sampling of this population. The findings would therefore appear to be suggestive, rather than conclusive.

It was found that some use has been made of charts, lantern slides, recordings, and especially motion pictures. Some 35 motion pictures, both silent and sound, were reported used in courses in clinical psychology. Generally speaking, films have been used most frequently in courses in developmental psychology, projective tests, personality therapy, and introduction to clinical medicine.

Some need was expressed for audio-visual aids to be used in courses in psychopathology, observational techniques, verbal intelligence tests, projective tests, guidance and counseling, and personality therapy.

On the basis of the above study, the following suggestions are made:

1. Information as to available audio-visual aids and their sources might well be furnished to instructors in courses in clinical psychology, so that the aids may be used when and if desired.

2. Evaluation of the effectiveness of available audio-visual aids in this field is greatly needed. This could be done in various ways: by individuals, by committees, by a central agency.

3. Needs of teachers of clinical psychology could be further studied. This might lead to revision of available films, as well as to production of new films.

BIBLIOGRAPHY

1. BETTIS, MOODY C., MALAMUD, DAVID I., & MALAMUD, RACHEL, F. Deepening a group's insight into human relations: A compilation of aids. *J. clin. Psychol.*, 5, 114–122.
2. GLAD, DONALD D. Roster of available films in psychiatry, mental hygiene and clinical psychology. Mimeographed, 1949. Colorado Psychopathic Hospital. Pp. 19.
3. KATZ, ELIAS. Audio-visual aids for mental hygiene and related areas. Reprinted from *Mental Hygiene News*, N. Y. State Hospitals Press, 1948. Pp. 6.
4. *Committee on Training in Clinical Psychology*. Recommended graduate training in clinical psychology. *Amer. Psychologist*, 1947, **2**, 539–558.
5. *Committee of Training in Clinical Psychology*. Clinical training facilities: 1948. *Amer. Psychologist*, 1948, **3**, 317–318.
6. *Psychological Cinema Register*. Pennsylvania State College, Pa., 1949.

Received August 31, 1949.

APPENDIX

FILMS USED BY INSTRUCTORS OF CLINICAL PSYCHOLOGY IN UNIVERSITIES ACCREDITED BY THE COMMITTEE ON TRAINING IN CLINICAL PSYCHOLOGY (JANUARY AND FEBRUARY 1949).

Aptitudes and Occupations. Coronet Films, Coronet Magazine, Chicago, Ill. 16 mm—sound, B&W, 20 min. Aptitude testing: mechanical, social, clerical, musical, artistic, scholastic, including practical applications.

Balloons: Aggressive and Destructive Games. New York University Film Library, New York 3, New York. 16 mm—sound, B&W, 20 min. Demonstration of a projective method for studying aggression in young children.

Behavior in Hypnotic Regression. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, 18 min. Behavior of a subject under hypnotic suggestion.

Care of the Sick and Injured by Hospital Corpsmen: The Neuropsychiatric Patient. Department of the Navy, Washington 25, D. C., 16 mm—sound, B&W, 10 min. Care of typical neuropsychiatric patients.

Child Development Series (16 films in series). Encyclopedia Britannica Films, Inc., 20 N. Wacker Drive, Chicago 6, Ill. Each film: 16 mm—sound, B&W, 10 min. A series of films by Dr. Arnold Gesell, Yale University, demonstrating patterns of child development.

Clinical Psychology and Hypnosis. Psychological Cinema Register, State College, Pa. 16 mm—silent, color, 15 min. A subject tested with Rorschach, TAT, Wechsler-Bellevue, etc.; hypnotic state and suggestion.

Conflict Situations in Childhood. University of Iowa, Iowa City, Ia. 16 mm—silent, B&W, 15 min. Kurt Lewin's method of studying behavior.

Convulsive and Allied Conditions. New York University Film Library, New York 3, New York. 16 mm—silent, B&W, 11 min. Demonstration of convulsive states.

Dynamics of an Experimental Neurosis: Its Development and Techniques. For Its Alleviation. Parts I, II, III, IV. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, I—20 min.; II—16 min.; III—19 min.; IV—20 min. Development of experimental neurosis in cats, and therapeutic measures.

Electroshock Therapy in Ambulatory Practice. Rutgers University Film Library, New

Brunswick, N. J. 16 mm—silent, color, 14 min. Indications, counter-indications and effects of electroschock therapy with ambulatory patients.

Emotional Health. McGraw-Hill Text Films, 330 W. 42 St., New York, N. Y. 16 mm—sound, B&W, 21 min. A college freshman undergoes psychiatric treatment for an emotional upset causing bodily symptoms.

Epidemic Encephalitis. New York University Film Library, New York 3, N. Y. 16 mm—silent, B&W, 18 min. Diagnostic and longitudinal studies of sequelae of epidemic encephalitis.

Experimentally Produced Neurotic Behavior in the Rat. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, 25 min. Rats develop various "neurotic" symptoms under experimental conditions.

The Feeble Minded. Psychological Cinema Register, State College, Pa. 16 mm—sound, B&W, 41 min. Diagnosis, classification, training of mental defectives.

Feeling of Hostility. National Film Board of Canada, Ottawa, Canada. 16 mm—sound, B&W, 31 min. Adjustment difficulties of a young woman who has achieved some measure of academic and professional success.

Feeling of Rejection. National Film Board of Canada, Ottawa, Canada. 16 mm—sound, B&W, 23 min. Adjustment difficulties of 23-year-old unmarried woman, and some psychiatric procedures for their handling.

Finger Painting: Children's Use of Plastic Material. New York University Film Library, New York 3, N. Y. 16 mm—silent, color, 20 min. Differences among 9 children in finger painting.

Frustration Play Techniques: Part I-Blocking Games. New York University Film Library, New York 3, N. Y. Different reactions of children in play-test situations involving conflict and frustration. 16 mm—sound, B&W, 21 min.

Frustration Play Techniques: Part II—Frustration and Hostility Games. New York University Film Library, New York 3, N. Y. 16 mm—sound, B&W, 20 min. Toys used to indicate degrees of frustration and release of hostility.

Grief. New York University Film Library, New York 3, N. Y. 16 mm. Prepared by Dr. Rene Spitz.

An Introduction to Clinical Neurology, Parts I, II, III, IV. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, I—19 min.; II—20 min.; III—17 min.; IV—17 min. Neurological tests, behavior disorders, functional disorders.

Let There Be Light. Office of the Surgeon General, Dept. of the Army, Washington 25, D. C. 16 mm—sound, B&W, 60 min. Methods of treatment of psychiatric disorders.

Meeting the Emotional Needs of Childhood. New York University Film Library, New York 3, N. Y. 16 mm—sound, B&W, 30 min. Demonstration of needs of child for independence and security.

Mental Defectives: Glandular Types, Mongolian and Cretin. Rutgers Film Library, Rutgers University, New Brunswick, N. J. 16 mm—silent, color, 32 min.

Narcosynthesis. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, 22 min. Four cases of narcosynthesis.

Paranoid State and Deterioration Following Head Injury. Psychological Cinema Register, State College, Pa. 16 mm—sound, B&W, 11 min. Interview with a 59 year old patient with paranoid state following head injury 10 years before.

Psychiatry in Action. British Information Service (British Consulate in all large cities). 16 mm—sound, B&W, 60 min. British hospital for rapid rehabilitation of war neuroses, both military and civilian.

Psychological Implications of Behavior During the Clinical Visit. New York University

Film Library, New York 3, N. Y. 16 mm—silent, B&W, 22 min. Individual differences at birth; contrasting maternal attitudes during nursing.

Psychoneurosis. New York University Film Library, New York 3, N. Y. 16 mm—silent, B&W, 16 min. Differential diagnosis of psychogenic and organic disease syndromes.

Shades of Gray. Office of the Surgeon General, Dept. of the Army, Washington 25, D. C. 16 mm—sound, B&W, 60 min. The neuropsychiatric problem in the Army.

Some Examples of Hypnotic Behavior. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, 11 min. Demonstration of hypnotic behavior reactions.

Symptoms of Schizophrenia. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, 18 min. Examples of symptoms in schizophrenia found in psychiatric wards of hospitals.

This Is Robert: A Study of Personality Growth in a Pre-School Child. New York University Film Library, New York 3, N. Y. 16 mm—sound, B&W, 75 min. A child at several age levels, demonstrating adjustment processes.

Treatment of Mental Disorders. Psychological Cinema Register, State College, Pa. 16 mm—silent, B&W, 20 min.

You and Your Child. U. S. Dept. of Agriculture, Washington, D. C. 16 mm—silent, B&W, 25 min. Aids in raising children in rural areas.

A NOTE ON LONDON'S HISTORICAL SURVEY OF PSYCHOLOGY IN THE SOVIET UNION[1]

GREGORY RAZRAN
Queens College[2]

For some 15 years some of the writer's colleagues have been urging him to write a book, a monograph, or at least a comprehensive *Bulletin* review of psychology in the Soviet Union. Somehow, the writer managed not to do so. Time and competing interests are some reasons for the failure. But an important consideration in the writer's mind has always been the difficulty of an objective appraisal of Soviet psychology and the necessity of becoming involved therewith in complex extra-empirical philosophies, criteria, and *Weltanschauungen.* Furthermore, the "First Principles" of contemporary Soviet psychology are so non-specific—really quite prior to and independent of any particular subject matter, and differing but little from Art to Zoology—that the writer has felt that any extensive treatment would duplicate much of what professional Soviet and Marxian theoreticians—and some non-Marxian philosophers, historians, and economists—have been doing for years. This of course does not mean that such a task should not have been undertaken, and London's efforts should be very much welcome. Yet his well-written article strengthens the writer's view of the difficulty with regard to objectivity and specificity.

Another difficulty under which London apparently labored is his inadequate access to original Russian publications.

1. A great deal of his discussion is based upon two textbooks—Rubinstein's *General Principles of Psychology* and Teplov's *Thirty Years of Soviet Psychology.* These two texts are quite important in themselves, but obviously are no substitutes for original reports and articles.
2. Of the 80 Russian references, 13 are articles in the *Large Soviet Encyclopedia,* hardly again an ultimate source material. In particular, the only references in London's bibliography to the leading Pre-Soviet and Early Soviet psychologist, Chelpanov, and to the outstanding psychological physiologist, Ukhtomskĭ—both of whom are discussed at some length in the survey—are such Encyclopedia articles. Several American libraries contain a good number of the original contributions of these two scientists.
3. Only one of the 14 references to Luria is Russian. True, Luria published considerably in foreign languages; still, his Russian repertoire surpasses his non-Russian one.

---

[1] London, I. L. A historical survey of psychology in the Soviet Union. *Psychol. Bull.,* 1949, **46**, 241–277.

[2] Prepared while the writer was a Fellow of the John Simon Guggenheim Memorial Foundation.

4. Of the 25 references to articles in Soviet psychological periodicals, 12 are from *Sovietskaya Pedagogika* [Soviet Pedagogy]. *Sovietskaya Pedagogika* began publication in 1937 and it devotes only a comparatively small portion of its space to psychology. More than a dozen periodicals wholly psychological in nature have appeared and reappeared in the Soviet Union in the last 30 years.[3]

Of London's omissions, the writer would like to note two: a school and a book. The school is that of Beritov and his many followers. Beritov is probably the most critical, most informed, and most psychological of all Soviet physiologists. The bulk and quality of his and his students' research and interpretations in neurphysiology, individual (conditioned) reflexes, and pattern and motor learning are outstanding by all standards. And lest the writer be considered biased, he would like to point, among other things, to E.B. Holt's *Animal Drive and the Learning Process* in which Beritov is quoted more than Pavlov, and to the very high regard of Beritov's work which K. S. Lashley expressed to the writer in a conversation a few years ago. The book that regretfully did not get into the survey is Bekhterev's *Psychology, Reflexology, and Marxism*. Apart from its intrinsic value, the book is a truly human document: von Bechterev, world-renowned neurologist and psychiatrist, liberal and constantly in trouble with the Tzar, pleading-proving by a mixture of naïveté and casuistry that reflexology *is* and psychology *is not* Marxian. Quite a variety of associated reflexes, exogenous and endogenous, seem to have participated in the performance.

Besides omissions, the writer would like to single out one inaccurate and misleading statement in London's survey. Discussing the disappearance of intuitive and speculative psychologies and the weakening of introspective psychology in the USSR, London states: "Instead, developments patterned after American behaviorism came strongly to the fore, especially in the guise of Bekhterev's reflexology" (p. 243). The statement is not only inaccurate technically because of the known fact that reflexology preceded behaviorism—noted by London himself elsewhere[4]—but also misleading in essence because (what London does not seem to be aware of) Bekhterev's reflexology remained little influenced by behaviorism even after the latter appeared on the scene. True, a good number of Soviet psychologists turned to behaviorism, and Watson's text underwent several editions. But this is of course an entirely different matter. There really were in the Soviet Union in the Twenties two separate "objective-materialistic sciences of behavior": an imported somewhat sophisticated one à la Watson and a cruder home-made one à la Bekhterev. The two had a lot in common. But there were also differences, and the differences might have been noted in the survey.

---

[3] It is also regrettable that London does not translate the titles of his Russian references; it reduces their serviceability to English readers.

[4] Bekhterev first formulated his principles of objective psychology in an article in 1904 (Obyektivnaya Psikhologiya i yeya Predmet [Objective psychology and its subject matter], *Vestnik Psikhol.*, 1904, 760–666; 721–737). The article was later published in part in French (La psychologie objective, *Rev. sci. Paris*, 1906, ser. 5, **6**, 335–337; 390–396). The date of the beginning of Bekhterev's system is thus 1904, and not 1907 as given by London and others; 1907 is the year of the publication of Bekhterev's book *Obyektivnava Psikhologiya* in which the earlier article was incorporated.

In general, however, London's achievement ratio with the material available to him is very high. Moreover, London apparently caught on to the spirit and the particularism of Soviet theoreticians, and, in the writer's opinion, his exposition would, with minor deviations, meet their approval. London is, for instance, right in saying that the writer's statements that dialectic psychology places "above all emphasis upon the social rather than the individual aspect of behavior" and that it "may probably be designated as some sort of Neo-Functionalism" are inexact dialectical renditions. However, the writer's excuse is that he wrote for an audience whose mastery of dialectics may well be subzero (the writer's own conditioned dialectics has undergone extinction, too), and that he used traditional psychological and not dialectical frames of reference and that within these frames his statements are correct and, he hopes, helpful.

The writer also hopes that in his second article, London will, together with a treatment of theoretical developments, make some quantitative comparison of the actual performance progress of Soviet psychology. Fernberger's count[5] of the number of Russian—among others—books and articles in the *Psychological Index* shows that while in 1929, 592 items, or 9.3% of the total, were Russian, in 1935 there were only 224 or 3.8%, Russian entries. And while it is true that only a portion of Russian material gets to be indexed and/or abstracted, the writer's perusal of the Russian bibliographical journals *Knizhnaya Letopis* (books published) and *Zhurnal'naya Letopis* (articles published) indicates a similar trend. Indeed, it is the writer's impression that even a sharper reduction in psychological publications took place in the USSR after 1935, and that psychology as a discipline occupies now a minor role in Soviet teaching and research.[6] It should not be too difficult to check the writer's impression by comparing publications, university courses, graduate degrees, and the like. It may well be that a deepening of dialectics entails a dwindling of data.

---

[5] *Amer. J. Psychol.*, 1936, **48**, 680–684.

[6] Would have been of even less significance but for the 100-year-old preoccupation of Russian physiologists with "higher" processes.

Finally, the writer is tempted to relate a somewhat amusing yet highly instructive episode in the history of Soviet psychology. It concerns the Yenchman movement or Yenchmanism. Yenchman was a Soviet behaviorist who attempted a synthesis of Marx and Watson by promulgating that "consciousness is an illusory-deceptive product of the capitalist class system and that it will wither away and disappear in the future socialist classless society." Furthermore, consistent with the dialectical principle of the unity of theory and practice and that of "changing, not merely understanding the world," Yenchman and his followers wanted to do something practical about consciousness. And seizing upon Bekhterev's speculations on the relation of consciousness to central inhibition and cortical resistance, the Yenchmanists set out to organize psychology clubs "wherein through a correct socialist peripheral uninhibited living consciousness could be more quickly liquidated." The clubs seem to have enjoyed popularity for a while but were disbanded as soon as Soviet psychology "regained consciousness." The liquidators (theories) were liquidated.

Received September 9, 1949.

A REPLY TO RAZRAN'S NOTE

IVAN D. LONDON
Tulane University

The writer's primary motivation in writing his historical survey of psychology in the Soviet Union was to present an objective account of the contemporary theses of dialectic psychology in a form intelligible to the non-dialectically oriented American reader. To realize this goal involved (1) translation from unfamiliar dialectical materialist jargon into comprehensible English—a task which turned out to be quite a chore—and (2) a historical approach, as the tenets of dialectic psychology are best understood against the background of the revolution and subsequent developments.

The writer has produced no apologia for the Soviet brand of psychology, but continues to entertain the not too radical notion that, before polemics or evaluation be indulged in, it is first necessary to have some factual basis preliminary to judgment. The work of three years, during which time the writer was also subject to the distraction of competing interests, convinced the writer that it should be possible to make a successful attempt at a preliminary account and elucidation of psychology in the Soviet Union both as regards to present tenets and history. He was particularly minded to do so because of the fact that no one else had, to date, seemed willing to take on the task of helping to fill in an undesirable gap in our knowledge of happenings elsewhere in the psychological world.

In view of the aims of the writer, he is unwilling, therefore, to concede the wholesale inadequacy of bibliographical material that Dr. Razran imputes. Rubinshteĭn's "Bases of General Psychology" is considerably more than a mere college textbook and the remaining bibliography sufficiently adequate for the writer's purposes.[1]

---

[1] The writer's complete bibliography extends considerably beyond that listed in his paper. References to Chelpanov and Ukhtomskiĭ (the former only briefly alluded to) are not limited to encyclopedia articles; e.g., **(7)** and **(125)** in the survey. Original articles, with the exception of **(122)**, were not consulted in this instance since the writer felt they were beyond the scope of his necessarily limited paper. The writer did not incorporate Beritov's school in his survey because, regardless of the high-level quality of its work, evidence seemed to indicate that its actual influence on the development of psychology in the Soviet Union was in no way commensurate with that of those schools discussed in the survey. Furthermore, footnotes 9 (p. 252) and 25 (p. 272) indicate that the writer was not unaware of the existence of a Russian behaviorism distinct from reflexology. The offending sentence to which Dr. Razran takes exception should have read, "Instead,

The writer is under no delusion that this first effort in the field is complete or definitive, subject to no correction, reinterpretation, or supplementation. On the contrary, the latter must in the very nature of the task be anticipated. It is for this reason that the writer welcomes Dr. Razran's observations and criticisms. The writer ventures the opinion, however, that his survey is in the main a reliable exposition and that its general outlines and trends will not suffer too much under subsequent reëvaluation and extension.

In conclusion the writer wishes to express his concurrence with Dr. Razran's estimate of the present minor role of psychology as a discipline in the USSR. In spite of the chauvinistic enthusiasm of its proponents and the fact that psychology is now a required subject in the secondary schools, the newly-formulated dialectic psychology remains as yet largely programmatic. It still begs for substantial realization in practice.[2]

---

developments to a degree paralleling or patterned after American behaviorism came strongly to the fore, . . . " (p. 243).

[2] *Editor's Note.* When asked if he wished to comment upon London's reply, Razran wrote: "Concerning Beritov, the writer would like to point out that Beritov is a member of The Soviet Academy of Sciences, a distinction shared by Pavlov and Bekhterev but by none of the two-score or so other names in London's article. (Rubinshteĭn is a Corresponding Member of the Academy.)" In reply to Razran's comment, London writes: "I don't see how Beritov's membership proves that he had an influence on the development of a dialectical psychology. The Russians don't think so. Bekhterev *did* have a decided influence in the definite sense that it was against him that the Russians directed their heavy artillery as a necessary preliminary to a psychology which was to be the antithesis of reflexology. Beritov, who was at first a "mechanist," later became a decided anti-reflexologist. The Russians reserved their ammunition for others. His contributions to a dialectic psychology were not important, particularly if we remember that the Russians distinguish sharply and somewhat differently between physiology and psychology."

Received September 30, 1949.

WHAT CONSTITUTES A GOOD TEXTBOOK IN SOCIAL PSYCHOLOGY?

DOROTHY G. PARK
Meredith College

By what criteria should one judge a textbook in social psychology? According to Littman's review[1] of Vaughan's *Social Psychology: The Science and Art of Living Together*, it would seem to be highly important that a text should be "academic" rather than popular, "analytic" rather than evaluative, and should stress "the logic of research" and "the nature of theory" rather than practical applications to the needs of society today.

Whether or not one agrees with these criteria probably depends upon one's fundamental point of view as to the aim of an elementary course in social psychology. Littman's aim would seem to be to make social scientists of his students, while Vaughan's, as stated in his preface, is "to show how we can apply to our everyday problems what we already know about human-beings-living-together, for better or worse, in modern society" p. vii).

A careful reading of Littman's review makes it clear that his criticisms of Vaughan's text are directed chiefly against the fundamental educational theory of the book, with which he finds his views in disagreement. Nowhere does he evaluate the text in terms of fulfillment of the author's avowed purpose, the only kind of really valid criticism. When he admits that "it will make interesting reading for beginning students," one wonders if, in the light of other criticisms, he finds this also objectionable.

In "progressive" circles it is considered a form of praise to label a text "definitely not academic," but Littman is not of that school. In the opinion of this writer the current textbook fare is so often starchily academic and difficult of digestion that frequent injections of vitamins are welcome. However, Littman objects to the "homely level" at which the volume is projected, to the ample stock of quotations from newspapers and the "better" magazines, to the use of "the language and notions of the layman" and to what he calls "popularized explanations." He seems to think that the recognized psychological procedure of finding the student where he is, meeting him on his own level and going on from there should be ignored in writing an introduction to social psychology. The failure to follow such a precept has been the chief drawback of many texts.

It may be instructive to consider Littman's one specific illustration of the "popularized explanations not much above the level of those it criticizes." He quotes from Vaughan's treatment of Le Bon's theory of the collective mind, adding, "it is disturbing to think that students might attempt to use such an explanation in meeting the complex problems of social life."[2] The reader has only to observe the footnote at the bottom of the page from which Littman quotes to discover (if he has not yet read the first chapter) that a logical discussion of Le Bon and the "group-mind fallacy" has already been presented on pp. 33–35. It is evident that the quoted sentences constitute a comment rather than a conclusion or an attempt at a "scientific account." In fact, the exposition of crowd behavior, with references to Le Bon, continues through the next eleven pages.

---

[1] LITTMAN, RICHARD A. Review of Wayland F. Vaughan's *Social psychology, The science and art of living together*. *Psychol. Bull.*, 1949, **46**, 314–316.

Littman insists not only that "the book is definitely not academic," but "is really of a type which the nineteenth century called 'moral philosophy', a free-ranging, evaluative . . . presentation," which "commits the error of advancing a specific set of values rather than providing some means of analyzing various normative proposals." In the face of the author's former achievements as a textbook writer of general psychology and his acknowledged use in this book "of the concepts and results of investigators having a distinctly systematic bias," Littman charges that "correctness by authority is substituted for a systematic account" and suggests that perhaps "the author believes that there is, in fact, no continuity between general or systematic psychology and social psychology."

It seems oddly inconsistent that Littman should object so strenuously to a philosophical bias, i.e., "a specific set of values," yet insist that the book falls short in not possessing a definite psychological bias. He cites no examples of the "personal hunches or guesses" which he claims "are often advanced as if they were *the*[3] answers to the problem." As a matter of fact, the author does indicate his views on such subjects as war, peace and democratic principles, but carefully follows his usual procedure of giving an historical background, presenting various viewpoints, using liberal quotes, and citing relevant experimental data from both psychology and anthropology. He seems to feel (and many would heartily agree with him) that any given psychological bias should be subordinated to the presentation of the social problem and its possible solutions in the light of what psychology has to offer.

Finally the reviewer objects to Vaughan's "emphasis on 'the here and now'." He feels that the author is too much influenced by his consciousness of social responsibility, with the result that he decides "to ignore the formal, conceptual processes of science" in favor of its practical applications. Littman would stress the "logic of research," "the nature of theory," and "insist that the student at least *try*[4] to acquire the conceptual tools of social psychology."

---

[2] *Op. cit.*, p. 315. Note: The page reference to Vaughan, cited by Littman is incorrect. It should read "p. 260."

[3] Italics are Littman's.

This appears to be the old controversy between pure and applied science. However, since this is a text for students it may well be that another criterion should be applied. The volume has grown out of the author's twenty years of experience in teaching social psychology to undergraduates and was written to satisfy the needs and interests of students. The vital issues then are: Can it meet these needs? Does it stimulate original thinking? Does it make psychology live in every-day life? The answer of the writer, who speaks from personal experience in the use of this text, is an unqualified affirmative.

[4] Italics are Littman's.

Received August 25, 1949.

BOOK REVIEWS

Krech, D. C. and Crutchfield, R. S., *Theory and problems of social psychology*. New York: McGraw-Hill, 1948. Pp. xv+639. $4.50.

This is a very important book which may well become a landmark in the history of social psychology. Its great importance lies in its systematic message—a brilliant and determined defense of the thesis that social psychology is, and must remain, part and parcel of the science of general psychology. At a time of unprecedented pressure for application of our knowledge to the cure of social ills, the social psychologist must more than ever be conscious of his roots in general, theoretical and experimental psychology. By bringing to bear the analytic methods and tools of his science on problems of social behavior he can contribute as a *psychologist* to the efforts of social science. The testing of general psychological theory in the area of social behavior will, in turn, broaden the horizons and enrich the scope of systematic psychology. Krech and Crutchfield's book is a demonstration of the fact that such a mature, systematic approach to social psychology is possible today. At the same time, this book shows that a systematic theoretical approach need not result in neglect of the urgent practical problems which fall into the province of the social psychologist. The authors strike a happy balance between problems of theory and practice, methodology and application.

Krech and Crutchfield are able to develop a coherent and consistent theory of social psychology because they have unfalteringly committed themselves to a particular systematic point of view. Theirs is clearly and thoroughly a field-theoretical approach. The strands of Gestalt theory of perception, Lewin's analysis of motivation and Tolman's theory of learning are clearly discernable in their exposition. This basic theoretical orientation remains intact through all the vicissitudes of application to problems of social behavior. Considering the present lack of vigorous theoretical debate in social psychology, such "dogmatism," whether one agrees with the position or not, seems preferable to eclecticism. The book shows how far one particular position can go, where its main strengths and weaknesses lie.

The plan of the book clearly reflects the authors' approach and intentions. The first part is devoted to a discussion of basic principles of motivation, perception and learning— a survey of the laws of cognition. The second part takes up the principles governing the formation and operation of those complex cognitive structures which are at the core of social behavior—beliefs, opinions and attitudes. The final part applies the generalizations and principles to the analysis of concrete social phenomena. We shall briefly describe each of these sections.

Part One. Basic Principles. In this section, Krech and Crutchfield develop their basic theme that the principles of motivation, perception and learning are the same for individual "non-social" behavior (e.g., behavior in the laboratory where interaction with other individuals is at a minimum) and social behavior. Much of the exposition is, therefore, devoted to general laws although the authors take every possible opportunity to reformulate general propositions so as to focus them sharply on problems of social behavior.

The authors begin with a discussion of the "dynamics of behavior," —the principles of motivation. The unit of analysis is "molar behavior" —unified, goal-directed sequences of acts. The dynamics of molar behavior must be understood from an analysis of the *immediate psychological field* at the time at which the behavior takes place. All motives are treated as contemporaneous, independently of their history in the lifetime of the individual. It is instabilities in this immediate field which produce "tensions" impelling behavior toward stabilizations of the field (tension reduction). The authors then go on to describe the manifestations of tensions in experience and in overt behavior, and the variety of actions by which individuals pursue their goals and attempt to achieve stable organization of their psychological field. The resulting picture of motivation is a far cry from a list of needs, instincts, or drives; rather, "needs and demands . . . are as variable as the tensions of the psychological field" (p. 72). Nor is simple "tension reduction" a guide to the prediction of behavior. Only infrequently does a psychological field return to its initial state of equilibrium. The trend of behavior change is typically toward progressively higher and more complex levels of organization. Only to the extent that there is uniformity of field conditions, is there uniformity of behavior, and the social psychologist is warned against assuming a core of invariant motives and response patterns which are independent of culture and social environment. His concern should rather be with the laws governing the functioning of fields—laws which are indeed universal.

Certain questions must be raised at this point. It would seem that "molar" behavior as a proper unit of motivational analysis needs sharper delineation. One man's molar is another man's molecular (thus Hull refers to his theory as a molar theory, which would certainly be considered molecular from Krech and Crutchfield's point of view). "Discrete unified episodes with a beginning and an end," to use the author's definition of molarity, may be of any degree of magnitude and complexity. One wonders whether it is necessary to posit a "proper" unit of motivational analysis and then to define the propriety, at least in part, in terms of the magnitude of the behavior sequence. In some cases, a highly restricted segment of behavior (a percept, a gesture, a verbal expression) may be illuminating to study in relation to social psychological variables.

The distinction which the authors draw between the study of the *immediate* psychological field and the genetic approach to motivation—

the contrast between historical and ahistorical analysis—may be too sharp. Granted that the field theorist does well to treat all motives as contemporaneous, the most useful and experimentally productive description of these motives, as they are effective *here and now*, may often demand historical analysis. To understand what the present motive is we may need to trace its history and development. Indeed, the authors recognize the importance of such historical analysis implicitly when they describe "the neural traces of past experience" as one of the major determinants of the immediate psychological field. To describe the past experiences whose traces are effective in the immediate psychological field perforce involves historical analysis of present tension systems. It would seem that the immediate dynamic problem and the genetic problem are not easily separated in the practice of motivational analysis.

The key feature of Krech and Crutchfield's treatment of motivation is their insistence that cognitive and motivational processes must not be treated in isolation from each other. It is through their effects on "*perception, cognition, and action*" that instabilities in the psychological field ("tensions") initiate goal-directed behavior. One of the crucial characteristics of motivated behavior is that it involves systematic cognitive changes. Analysis of motivation cannot long proceed without description of cognitive changes, nor can analysis of the cognitive processes be taken out of the context of the goals which the individual pursues. Thus Krech and Crutchfield's position attempts to do away with the question, "what is the influence of motivation on perception?" and equally would render obsolete the question, "what is the influence of cognition on motivation?" All psychological fields are characterized by cognitive structure and varying degrees of instability (tension). It is selective emphasis upon one or the other aspect of the psychological field which leads us to retain the distinction between cognitive and motivational processes.

Thus the analysis of motivation leads the authors naturally into the problems of perceptual organization and reorganization, to the study of the cognitive characteristics of the psychological field. The central significance of perception to a systematic theory of social psychology is by now clear; all action, including social action, is determined by the ways in which the environment (including other individuals) is perceived and interpreted by the individual. The systematic social psychologist faces two major tasks: first, he must learn to be a good phenomenologist, to describe the social environment as it is perceived by the individual. Secondly, he must participate in the formulation and testing of general laws of perceptual organization, for the same basic set of laws governs the perception of social and non-social environments (if there be such a thing as an entirely non-social environment). Krech and Crutchfield make an attempt to carry out both these objectives. They give us a series of propositions which they believe embody the basic principles of perceptual organization and reorganization, drawing their illustrations largely from the field of social behavior but continually pointing to parallels in "simple" visual perception.

The authors distinguish two types of determinants of perceptual organization: "structural" and "functional" factors. The structural factors represent the capacities and response characteristics of the nervous system; they set, as it were, the limits within which a perceptual response to physical stimulation can vary. The functional factors represent the needs, goals, and past experiences of the organism; it is these factors which are largely responsible for the variations in perceptual responses from individual to individual, from group to group, and from culture to culture. Consistently with their analysis in terms of total psychological fields, the authors insist that every perceptual organization is jointly determined by both types of factors. Functional and structural factors are necessary abstractions for the sake of experimental manipulation and analysis. Within this theoretical framework, the authors then present a set of specific propositions stressing the organizational properties of the perceptual field (very much in the Gestalt tradition), its functional selectivity determined by needs and mental sets, the differentiation and integration of perceptual fields, and the operation of such laws as similarity and proximity (again along the lines of Gestalt theory).

In presenting their propositions, Krech and Crutchfield specifically state that they do not propose to prove their validity by experimental evidence nor do they wish to evaluate the relative contribution of structural and functional factors. One wonders whether the cogency of their argument does not suffer from this omission. The status of the "propositions" is not entirely clear—in what measure they are empirical generalizations, in what measure axiomatic statements, and in what measure statements of theoretical interpretation. Of course, we should not expect a full-scale exposition of perceptual theory in this part of the book. It might have been helpful, however, to separate axiom, empirical fact and theory, all of which may be discerned in the propositions as stated. And the future researcher would have been aided by an explicit discussion of the ways in which structural and functional factors can be teased out experimentally.

Cognitive *reorganization* is the avenue by which Krech and Crutchfield approach the problems of behavior change—the problems of learning, retention, and emergence of new needs. There is a radical departure here from conventional treatments of social learning. What is usually described as "learning" is treated as a special case of cognitive reorganization in time. To understand the determinants of behavior change, argue Krech and Crutchfield, we must understand the determinants of cognitive reorganization. The final chapter of this section, then, presents a series of propositions stating the conditions—"structural and functional"—under which such cognitive reorganization occurs and the basic dimensions of change. The treatment is strongly influenced by Gestalt theory and by Lewin's position. Thus, "the ease and rapidity of cognitive reorganization process is a function of the differentiation, isolation, and rigidity of the original structure" (p. 135), and cognitive reorganization will continue as long as there is "blockage to the attainment of a goal" (p. 112). The treatment of memory favors Gestalt interpretation and Bartlett's schema theory.

To the social psychologist, one of the major problems of behavior change has been the process of "socialization." He has been searching for laws which govern the acquisition by the child of the ways of his environment. In the opinion of this reviewer, Krech and Crutchfield have made a real contribution to the teaching of social psychology by not treating socialization as a "separatist" problem requiring the formulation of special laws of social motivation and social learning. The psychologist who has confidence in his principles of learning and motivation must be willing to let these principles serve him in what is perhaps their most vital area of application—the process of socialization. If the principles should prove inadequate, then the remedy must be sought in a revision of the principles, not in an ad hoc theory of socialization. There is one possible weakness, however, in Krech and Crutchfield's treatment of the learning process: they have not concerned themselves sufficiently with the acquisition of *skills*, with the ways in which cognitive reorganization is translated into motor action. In this connection, it would seem that the law of effect and the problem of reinforcement (and their possible reinterpretation in terms of cognitive theory) have been given too short shrift by the authors.

Part Two: Social Processes. This section is devoted primarily to the study of those "enduring mental organizations" which are postulated in the analysis and prediction of social behavior—beliefs and attitudes. The distinction which Krech and Crutchfield make between these two concepts is of interest. A belief is defined as an "enduring organization of perceptions and cognitions about some aspect of the individual's world" (p. 150). Such beliefs may range along a scale of verifiability, from knowledge to opinion to faith. An attitude is "an enduring organization of motivational, emotional, perceptual and cognitive processes with respect to some aspect of the individual's world" (p. 152). The basic difference appears to be that beliefs are relatively neutral whereas attitudes embody impulses toward motivated action, i.e., the objects of attitudes are perceived as goals.

Is it useful to perpetuate such a distinction in terms of Krech and Crutchfield's own theoretical position? They have insisted throughout that every psychological field has both motivational properties (degrees of instability or tension) and cognitive properties (selective perceptual organization, degree of differentiation, etc.). The dividing line between belief, knowledge, faith, opinion and attitude must necessarily be thin, and in any given concrete situation it might be difficult indeed to state which of these postulated cognitive structures was in operation. It would seem more consistent with the authors' general argument to say that cognitive structures vary along such dimensions as permanence and degree of "dynamic pressure to action." The emphasis would then be on the measurement of these dimensions in relation to concrete behavior rather than on the measurement of attitudes and beliefs which overlap in uncertain ways. There may be more at issue here than a question of terminology. The theoretical problem which merits further discussion is whether in a systematic approach like that of Krech and Crutchfield such concepts as attitude have not outlived their usefulness. Their theory should be free of the necessity to provide a niche for time-honored concepts that do not fit organically into the system.

Following the definition of beliefs and attitudes there is a thorough and well documented discussion of the kinds of attitudes which have been studied, of their dimensions of variation and "functional significance"—the ways in which beliefs and attitudes help to render the environment more coherent and meaningful and are instrumental in the solution of daily problems. The authors raise the thorny problem of what the factors are—environmental and motivational—which determine the development, maintenance and change of attitudes, steering a careful and critical course between extreme culture-centered interpretations on the one hand and extreme personality-centered views on the other. Beliefs and attitudes are developed selectively by the individual within the limits of a range set by the culture. Finally, the authors give practical application to their analysis by developing a programmatic guide for those who face the formidable task of changing beliefs and attitudes.

The sections on measurement of beliefs and attitudes, including scaling techniques and public opinion research, are remarkable for their thoroughness, clarity, and interesting presentation. The authors deserve thanks for teaching the methodology of opinion measurement in a way which combines technical sophistication, a wide range of application and great sensitivity to the critical research needs in this area.

After the chapters on measurement of opinion comes an analysis of the technique which is so frequently employed in the hope of changing opinion—persuasion by propaganda. The main systematic point of this analysis is the authors' insistence that neither propaganda nor its handmaiden, suggestion, is in any sense a unique force or a phenomenon requiring special laws of its own. What we call "suggestion"—whether in the laboratory or in a propaganda campaign—is a manifestation of the same general laws which govern all cognitive organization and reorganization.

In the final chapters of this section the "structure and function" of social groups are discussed. The authors reiterate their position that there are no social phenomena—including group behavior—which should not ultimately be reducible to explanation in terms of the general psychology of the individual. In their own words, "the study of groups as a whole can reveal nothing new beyond what is given by a synthesis of all the data pertaining to each of the group members" (p. 306 f.). For pragmatic reasons, however, Krech and Crutchfield deem it advisable to switch at this point to analysis at the group level. Their discussion of group structure deals with such problems as the effects of group size, the distribution of roles within the group, and such methods for measuring group structure as sociometry. The functional analysis of groups is concerned with the needs and demands which groups meet and create in their members. The patterning of beliefs and actions in groups, and problems of morale and leadership, round out this discussion.

With the pragmatic wisdom of including such an analysis of the behavior of groups one must agree. The systematic question remains, however, to what extent such an analysis is an organic part of the theory of social psychology whose programmatic outlines Krech and Crutchfield have sketched for us so cogently in the first part of the book. In this question there is the seed of a conflict for social psychologists: whether to put foremost the task of psychological theory construction which may, at least temporarily, narrow the scope of their subject matter, or whether to extend the borders of the discipline at the expense of systematic cohesiveness.

Part Three. Applications. This section is devoted to an application of the principles and concepts developed in the first two parts of the book to a number of urgent social problems of the day: race prejudice, industrial conflict and international tension.

In their discussion of these topics the authors' main emphasis is on formulation of the problems in terms of individual psychology. They are fully aware of the importance of setting these major social disorders in their proper social, political, and historical context and indeed devote part of their discussion to providing this context. They focus, however, on a psychologist's view and the remedies and reforms for which they plead are couched in psychological terms. They ask questions such as these: what are the major needs and demands which are satisfied by prejudice or which give rise to industrial tensions and international unrest? What cognitive organizations and reorganizations (beliefs and attitudes) are crucial in these areas of social behavior? What are the environmental supports which maintain prejudice and tension? How can motivational control and cognitive training help to cure these ills?

It is no reflection on the authors' skill that this last section remains to a considerable extent programmatic. It becomes clear how a systematic theory such as that of Krech and Crutchfield will go about attacking the formidable task of application, but the major task of formulating specific hypotheses in these areas and testing them in the teeth of tremendous practical difficulties is still ahead. The important point to stress is that the authors' concern with theory leads them to clearer and more specific formulations of applied problems than might otherwise have been possible.

In conclusion, thanks are due to the authors not only for a challenging and stimulating treatise but also for the felicity of their style, their careful and comprehensive documentation, and their flair for making their subject matter alive and interesting. It is a good book from which to teach, and a good book to argue about and to ponder.

Harvard University. LEO POSTMAN.

BRITT, STEUART H. *Social psychology of modern life.* (Rev. Ed.) New York: Rinehart, 1949. Pp. xvi+703. $4.50.

This book, while rather completely rewritten, still has most of the characteristics of the first edition, including both its virtues and its defects. The level of writing remains semi-popular, which is defensible in view of Britt's stated purpose, viz., to write for beginning students and interested laymen. The range of topics continues wide, from subliminal stimulation to nationalism and war. The choice of illustrations is eclectic and catholic, as laboratory studies rub elbows with anthropological observation and tabloid newspaper clippings.

As one reads this book, he finds it congruent with the author's post as Director of Research for a large advertising agency. Britt has a shrewd eye for effective illustrations (p. 70) and amusing quotations (p. 190). There is a great deal of visual material, although some of it seems to relate only to minor points, hardly worth the space. It is a pleasure to note that the author's economic connections have not blinded the critical eye with which he still inspects our social and economic institutions.

On the debit side we first become aware of a certain thin quality to the book. The theoretical analysis of some problems seems better than in the first edition, but on the whole it is not such as to be valuable in any but an introductory psychology class. Thus, although the author devotes a good deal of space to race prejudice, there seems little effort to systematize the theoretical discussion. This may have been deliberate, as most of the necessary elements are casually mentioned. There is, however, a superfluity of illustrative material, to the point of confusing the reader in some cases.

Less easy to excuse is some considerable carelessness in the use of technical terms. Thus the use of "differentiation" (p. 100, p. 206) does not conform to that sponsored by Allport, from whom it seems to be borrowed, nor to that of Pavlov, to whom the discussion refers. Similarly, one wonders about the sharply limited space allotted to motivation (6 pages shown in index, and not much more woven into the text) and perception (listed under "Observation" in index, illustrated but not systematically treated in various chapters). Frame of reference gets only one mention, as does the concept of behavior norms.

Another decided inadequacy is in the field of methodology. It does not seem plausible that attitude measurement can be dismissed in two pages, with a casual mention of the opinion census, the "*a priori*" scale (e.g., Bogardus) and the psychophysical scale (Thurstone). Experiments in social psychology are discussed more to show their futility than to indicate their possibilities; this may be correlated with the Index, which gives six references to Kurt Lewin, and fifteen to William James.

On the whole, Britt has given attention to the real and significant problems of social psychology. With the exception of industrial conflict, he has touched on most of our current problems, shown their irrational components, suggested that psychology can make a contribution. His final chapter, "Knowledge for What?" pulls no punches. And yet one is likely to be left with the feeling, "Social Psychology for What?" Britt has not come to grips with the theoretical problems which must be at least envisaged if the reader is to have a clear orientation in this maze of complexities.

ROSS STAGNER.

University of Illinois.

LAPIERE, RICHARD T., & FARNSWORTH, PAUL R., *Social psychology.* (3rd Ed.) New York: McGraw-Hill, 1949. Pp. xiii+626.

The first edition of LaPiere and Farnsworth appeared in 1936, when psychologists generally were at war with sociologists over the field of social psychology. The book was proof that collaboration between representatives of these two disciplines could result in good to all concerned, especially to students in social psychology. The second and third editions have done even more to enlarge the horizons of students who might otherwise have been guided more strictly into one or the other of the two principal approaches to the study of social phenomena. The ample documentation, the appendices of running comments, and bibliography have proved useful to all who have wanted to read beyond the text.

As earlier editions have made this a standard textbook of known merit, it would be superfluous to describe the contents of the third edition in detail. Part I, *The Nature of the Individual and of Society*, retains essentially the same flavor that it had in the first and second editions. To some, it may seem unfortunate that the dichotomous concepts of *individual* and *society* are retained in the subtitle and throughout the text. The authors continue to pay their respects to the concept of learning early in the text and then shift to the concept of socialization for purposes of explanation and description, as if *learning* and *socialization* were separate and distinct phenomena.

Part II, *The Process of Socialization*, contains a chapter, largely new, entitled *Socialization*. In the second edition the chapters on gesture and speech were combined; the third edition is like the first in devoting a chapter to each of these topics. Part III, *The Human Personality*, has a new chapter on personality and one on personality and social controls. Now so many departments of psychology offer separate courses in social psychology and in the psychology of personality that one might expect the discussion of personality to be mimimized in a textbook on social psychology; however, sociologists find it difficult to omit this interesting topic. Part IV, *Personality and Social Adjustment*, and Part V, *Social Interaction*, though rewritten in part, do not deviate greatly from earlier editions.

The systematic orientation of the book is perhaps to be found in the statement in the new *Introduction* that "the behavior of man is largely a product of the behavior of other men, known collectively and abstractly as 'society'." The recent emphasis on perception and learning in social psychology is scarcely represented in this work and it has not greatly affected the systematic position of the authors. For them, social psychology seems to remain partly sociology and partly psychology, with the concepts of sociology predominating, as is consistent with the history of social psychology. Yet the collaboration between the authors is so wholehearted and harmonious that speculation as to which parts were contributed by the sociologist, and which by the psychologist, would be idle.

Williams College. ARTHUR JENNESS.

POLLAK, OTTO. *Social adjustment in old age.* New York: Social Science Research Council, Bulletin 59, 1948. Pp. xi+199.

This work is intended as a guide to planning research projects on the social adjustment of old people. For practical purposes old age is considered to begin at sixty years, but it is emphasized that, functionally regarded, it varies considerably above and below this point in time of onset.

The author proposes, as a frame of reference for research, a concept of the problem of old age which envisages it in terms of adjustment. The well-adjusted person is regarded as one who is able to satisfy his needs, quickly and adequately, as they arise. This individual approach is then broadened into a social one, in which concepts of anthropology and sociology are added to those of psychology.

The remainder of the book is largely an amplification of the approach just outlined. Many different phases of the problem of aging are considered, and valuable information is given about how such problems have been and might be attacked. In a broader sense, the two most valuable contributions are probably the frame of reference for study which is offered, and an emphasis upon the great importance of a full consideration of demographic factors as a basis for research on the aged.

The book is disarming in its objectivity and in its transcendance of the limits of separate fields of inquiry. It is free from any tendency to limit or to direct the course of future research. It leaves the individual researcher free to follow his bent. It does, however, provide for his benefit considerable helpful wisdom derived from a pooling of the thought and experience of many investigators.

To the reviewer the chief limitation in the author's work is the slight attention given to the possible study of well-adjusted old people. Of course, as in most fields dealing with human adjustment, the case studies which have been made are of the maladjusted rather than of the well-adjusted. This is probably inevitable when the emphasis is on cure rather than on prevention. Might it not be well, however, to take also the long-term view, and to make careful studies of the factors which have been enabled some people, at various social and economic levels, to grow naturally into a happy old age?

JOHN T. METCALF.

University of Vermont.

SCHAEFER-SIMMERN, HENRY. *The unfolding of artistic ability.* Berkeley: Univ. of California Press, 1948. Pp. xiii+201. $5.00

This book presents a notable contribution to a theory of artistic creativity, a theory that is supported by facts and, particularly, by educational success. The theory is based on the views of the German leading art teacher and philosopher Gustav Britsch; according to him, children's drawings not yet distorted by external methods of teaching display a lawful development toward structurally ordered representation which, in essence, is inherent in every work of art. This structurization follows general laws of perception, and it is these principles of configuration which enable the artist to express his ideas within the framework of a "visual system," as it were.

All normal children, according to the author, display an inner drive for pictorial cognition which, with the majority of individuals, declines during the period of growth as a consequence of the emphasis of our culture and educational philosophy upon abstract knowledge. Schaefer-Simmern insists that there are dormant creative activities in everyone which can be made to unfold, due to the inner genetic logic of artistic expression. A consequence of education in artistic expression is the cultivation of disciplined feeling, perceiving, thinking, and acting; the unfolding of artistic activity assists thus in forming a balanced personality.

In order to prove the validity of these hypotheses, educational experiments with persons of wide variety in mental capacity and cultural background were performed. Particularly striking are the author's successes with mental defectives; they clearly demonstrate that many feebleminded individuals can be raised to a level of organized activity far beyond the stage indicated by the intelligence score of these persons.

Schaefer-Simmern's work illustrated by many impressive drawings is a significant pioneer venture in the field of art education. For experimentally-minded psychologists, the demonstration of an apparently universal trend toward expressing visual experience through organized form and the effect of this expression on mental maturation is a challenge to further study these processes by means of scientifically controlled methods.

HEINZ WERNER.

Clark University.

ANDREWS, T. G. (Ed.) *Methods of psychology*. New York: John Wiley, 1948. Pp. xiv+716. $5.00

Methods of Psychology is presented as a partial solution to the problem of teaching psychological methods to the undergraduate student of psychology. It is meant to serve as a textbook in a course in methods or as a supplement to more "factual" books in other courses. The volume contains 22 chapters, each written by a different author and each covering one area of research. A short bibliography is given at the end of each chapter. The material covered includes such diverse fields as sensory, physiological, abnormal, and social psychology, conditioning, learning and testing. The topics discussed are indicated by the following list of chapter titles (the names of contributing authors are given in parentheses): An Introduction to Psychological Methodology (T. G. Andrews), Conditioning and Motor Learning (W. N. Kellogg), Studying Memory and Transfer (C. E. Buxton), Studying Human Thinking (E. Heidbreder), Psychophysical Methods (L. L. Thurstone), Studying Perceptual Phenomena (J. J. Gibson), Studying Vision (S. H. Bartley), Studying Hearing (E. G. Wever), Studying the Skin Senses (W. L. Jenkins), Studying the Senses of Taste and Smell (C. Pfaffmann), Studying Proprioception (W. D. Neff), Studying Animal Behavior (H. F. Harlow), Motivation, Feeling, and Emotion (P. T. Young), Methods of Measuring and Recording Action (R. C. Davis), Studying Neuropsychology and Bodily Functions (D. B. Lindsley), Studying Motor Functions and Efficiency (A. G. Bills), Investigating and Appraising Intelligence and Other Aptitudes (H. S. Conrad), Investigating and Appraising Personality (S. Rosenzweig), Methods and Techniques in Clinical Psychology (A. W. Brown), Objective Studies of Disordered Persons (J. Zubin), Methods of Studying the Behavior and Development of Young Children (H. L. Koch), and Studying Social Behavior (T. M. Newcomb).

The term method is a broad and ambiguous term. It is probably true that the usual distinctions between method, data, and theory do not permit of explicit and rigorous expression and one could discuss at great length the feasibility of writing about one without also writing about the other two. That it is possible or profitable to discuss what investigators do without discussing what they find may be seriously doubted by most workers. Nevertheless, it is probably equally true that, at various stages in the development of a science, gross classifications of problems and subject matter are possible and useful, at least for instructional purposes. Provided we do not attempt to carry the analysis too far and consider the implicit assumptions and theory involved, it should be possible, for example, to discuss apparatus used by the psychologist without considering all of the rational bases for its use.

Even if we grant that a differential emphasis of problems is possible, the task remains to specify what is to be included under each gross heading—in the present case, method. It is at this point that the reviewer feels that *Methods of Psychology* has failed. The decision to minimize "involvement in purely historical and theoretical controversies" and to hold "the sections on results. . . . to a rather minor role" has apparently not provided an adequate enough specification for the contributors of what should be discussed. The specification of a few things to minimize has done little to decrease the ambiguity of what has to be maximized. To discuss methods in a given area may, for one person, mean to outline and discuss the variables that must be controlled in an investigation; for another person, it may mean the next step, to discuss the techniques and apparatus by which the variables are controlled; for another person, it may mean neither of these, but it may imply a logical analysis of certain experimental procedures; for others, it may have different meanings. The contributing authors of *Methods of Psychology* provide a sample of the variety of meanings that attach to the word method. In fact, one characteristic of the book is the diversity of conception of the subject matter to be presented. Some of the differences between contributors in the treatment of method may arise from the variety of areas of psychology considered. In some areas of psychology questions of technique of controlling variables are more disturbing than are the prior questions of what variables to control. For example, if we discuss wave length of a visual stimulus as a variable, the answer to the question of how one controls this variable may not be obvious to the student; whereas, if we discuss time interval between trials in a learning situation as a variable, the question of how to space trials will seem trivial to the student. All of the differences between chapters do not arise from this source, however. In the field of learning for example, where problems of apparatus are not overwhelming, we find a representative treatment of techniques of studying learning; while in the field of sensory psychology, where problems of stimulus control are more in evidence, we find a chapter on vision that has little to say about apparatus.

The multiformity of presentation makes an evaluation of the book difficult. An evaluation of each chapter in terms of what the author is trying to do yields a result different from an evaluation of the book as a whole. Some of the chapters, for example, Lindsley's chapter, *Studying Neuropsychology and Bodily Functions*, are good rounded reviews of how problems are studied in specific fields. Other chapters are examples of particular approaches to the discussion of method. Harlow's chapter, *Studying Animal Learning*, gives a straightforward presentation of method from the point of apparatus employed. It presents figures and brief discussions of many of the simple, but numerous, devices for studying behavior in lower animals. Buxton's chapter, *Studying Memory and Transfer*, gives an explicit account of method in terms of variables that must be controlled. Zubin's chapter, *Objective Studies of Disordered Persons*, considers method in that area by giving a detailed discussion of a small number of experiments. Most of the chapters on the senses are good, although the chapter on vision is weak. With some qualification, the chapters on conditioning and learning are adequate for a second-year course. The chapter on perception is weak. Davis' seems out of place. Perhaps this is because there are not more chapters like it. This chapter considers methods of amplification in recording, ways of taking permanent records of responses, etc. The reviewer expected more discussion of apparatus and technical problems of this sort. As it stands, however, the chapter does not seem to fit into the rest of the book.

Psychology texts continue to contain passages that sound like psychology but whose meaning is elusive, and the present volume is no exception. For example, in considering the problem of the "constancy" phenomena we are told (p. 180) that "If retinal shape does not correspond with seen shape, neither the basic assumptions nor the methods of psychophysics are applicable to the problem." This is a straightforward statement although it places an unjustified restriction on the psychophysical methods. Any doubts the reader may have concerning this restriction of psychophysical methods will not be removed by the list of alternative methods indicated in the next sentence. Here we are told that the problem "will have to be studied in terms of mental transformation, interpretation, intuition, or at best, a process of spontaneous sensory organization." This quotation is not meant to be a representative sample from the book, which, in general, has a small number of such literary excursions. The quotation does indicate one kind of fault that may be found and a confusion that will continue to plague the student of psychology.

One could consider each chapter individually and attempt to arrive at an average for the book as a whole, but the question remains of how these and other chapters combine to give the reader an integrated picture of the methods of psychology. Here the answer is less positive. The great diversity of treatment of method given in the work has some advantages, but the reviewer feels that much would be gained by giving a more uniform treatment of at least one conception of method. (It may be that a more restrictive title would then be appropriate.) Although many of the individual discussions are good, the presentation of apparatus in one area, controlling variables in another and possibly neither of these in a third area destroys whatever continuity may be possible with such an expansive subject matter. If the space in a book of this size does not allow a discussion of the many ways in which the term method is used, the reviewer would favor a systematic coverage of one view of method, i.e., a discussion of controlling variables for all areas of psychology treated or a discussion of apparatus and equipment for all areas, etc. This opinion may not be generally shared. The treatment found in Andrews' book has the advantage of showing the many ramifications of the answers to questions of method.

In terms of difficulty, the book is directed toward the second-year student and most of the chapters are written appropriately toward this goal. The chapter on psychophysical methods, although a good chapter, may be difficult for the average student. Several other chapters may be difficult because some writers have, in an effort to include all phenomena, classified and categorized beyond the stage where simplification results. At an early stage of development of an area of research, exhaustive classifications lead to as many classes as there are phenomena and the advantages for the student are not always obvious.

Finally, it is likely that the undergraduate student's appreciation of the place of method in a science and his general orientation to methodological problems would be enhanced by more frequent reference to the functions which methods serve and the relations that exist between method, data and theory. In particular, more frequent reference to the way in which data and theory influence method would probably serve the useful end of making the student appreciate the significance of particular procedures. The emphasis on the function of particular experimental methods rather than on the methods *per se* would show more clearly the elegance of certain procedures. It should also reduce the tendency to study problems with the same procedures that have always been used and to ask only those questions which may be answered by existing techniques. Greater emphasis on the interaction of method, data, and theory would allow the student to appreciate the fact that, in the final analysis, the distinctions are useful mainly for instructional purposes and for textbook and article organization.

CONRAD G. MUELLER.

Columbia University.

SNYGG, DONALD, & COMBS, ARTHUR W. *Individual behavior: A new frame of reference for psychology.* New York: Harper, 1949. Pp. ix+386. $3.50.

The last half century has produced many theoretical frameworks within which to organize the empirical facts of psychology. Each in turn has demanded attention; each has had varying degrees of acceptance; each has in some degree waned while at the same time it has contributed a permanent addition to the total body of psychological science. From the mentalistic positions, which were a part of philosophical discussion for two millenia, the trend in theories was toward greater objectivity—psychology was to deal with concepts as free from metaphysics as were those of physical science. The basic datum of psychology was behavior as observed by someone other than the organism behaving. The authors of the present volume call this the "external approach"; they consider it entirely unsatisfactory because, they claim, it permits prediction of behavior only on a normative basis—it is unable to predict exactly the behavior of the specific individual. Such individual prediction the authors appear to hold is the *sine qua non* of psychology, especially as it is applied to human affairs.

In the present volume the authors present the phenomenological or "personal" approach to behavior as a frame of reference in lieu of the "external approach." In spite of the book's sub-title, "A new frame of reference for psychology," the contents, and the authors' direct statement (p. 206), indicate that this point of view has been discussed by a number of psychologists (p. 11). The real contribution of this volume is its essay at an integrated and detailed development of the point of view. Part I, with 8 chapters, presents a logical development of the systematic position, while the one chapter of Part III contrasts the phenomenological position to mentalistic, neurological, and other explanatory principles as scientific methodology. Part II, with 6 chapters, illustrates the place of the personal approach in application to social psychology, education, diagnosis, and psychotherapy. It is neither necessary nor possible to examine all of the book's details. If one grants the adequacy of the basic postulate and its corollaries, one must accept the development therefrom. While minor apparent contradictions may be found they do no violence to the logical development. In the discussion of application this reviewer had occasional impressions that the samples were being "slanted" and that consistency was somewhat forced. A serious critique, however, must consider the basic postulate.

In the words of the authors, "*All behavior, without exception, is completely determined by and pertinent to the phenomenal field of the behaving organism.*" Two terms in this basic postulate need defining. The phenomenal field is "the entire universe, including himself, as it is experienced by the individual at the instant of action." Thus behavior is not related to situations as such but to the behaver's perception of the situation. There can be no quarrel with this principle, but it is not at all new, the essence having been presented by J. R. Kantor (who is completely ignored by the authors) in the concept of "stimulus function" in his entirely objective system published in 1924–26.

The second term which merits attention is "completely determined." While not explicit in the formal statement of the basic postulate, it is evident from many other statements in the book that "completely determined" is equivalent to "caused," in the full metaphysical connotations of the latter word. In fact, in their criticism of the "external approach" the authors express doubt of its scientific status because of the avoidance of the concept of "causal relation" (p. 7). The insistence on the importance of metaphysical cause in this exposition of the "personal approach" appears, to this reviewer at least, to cast some doubts on the scientific adequacy of the authors' point of view.

The phenomenal field is the determiner of behavior, in a causal not in a statistical sense (p. 15). Individual behavior can be predicted from knowledge of the behaver's phenomenal field (p. 45). Such knowledge cannot be secured from direct observation, but must be inferred from observed behavior (p. 21). These three sentences, I believe, do no injustice to the authors' discussion of prediction. (A check of each reference under the entry "prediction" in the index failed to reveal any clear-cut explanation of the process.) It would appear that in their effort to secure exact individual prediction the authors prepare an unnecessarily cumbersome machinery. If the phenomenal field can only be inferred from observed behavior it hardly seems necessary to introduce the entity at all. In short, prediction is made on the basis of past behavior, and this is exactly what an "objective" system does. Similarly, on page 29 the authors renounce past experience together with mind and the nervous system as explanatory principles. Yet on page 83 they say that the phenomenal self (a part of the total phenomenal field) is a result of experience. This being true, it is difficult to see why the "phenomenal field" is a necessary entity. The most serious question to be asked is "Why is it necessary to posit an inferential entity, which is analogous to mind, or even the nervous system (in spite of the authors denials), when the same facts can be explained without violating the Law of Parsimony?"

In one sense this review is not entirely fair to the volume being reviewed. Because, as a systematic position, the whole discussion rests upon the basic postulate, my primary attention has been directed to it. Even if one cannot accept the postulate the work has real significance. Clinical or other applied psychologists especially should study the proposals carefully. The details of development of the "personal frame of reference" have real prognostic value, even though question may be raised as to the theoretical adequacy of the system.

C. M. Louttit.

University of Illinois.

APPICCIAFUOCO, ROMOLO. *Sommario di Psicologia.* (*Review of psychology.*) Rome: Orsa Maggiore, 1949. Pp. 243.

Of the forty-two nations that met in Geneva in 1937 in the International Bureau of Education, only Italy did not agree to include psychology in the required curriculum of teachers. This fact emphasizes the inferior position that Italy holds today in psychology, despite such eminent men as Giuseppe Sergi, Gregorio Rossolimo, Giorgio Zunini, and Sancte de Sanctis.

The appearance of this book in 1949 is a distinctive surprise and may be the harbinger of a new day in Italian psychology. If translated, it would compare favorably with the standard introductory volumes available to the American student. Although the emphasis is chiefly continental, American psychologists such as Giovanni Dewey, Giuseppe Jastrow, Enrico Murray and Luigi Terman are well represented.

The contents of the volume are broken into two sections, theoretical and applied, each consisting of ten chapters. Although direct references are not made, it is evident that the author is well up on the latest thinking in the various areas, and is able to integrate European and American investigations in a smooth and effective manner. However, in many cases European authors are emphasized over the more important American writers. As an example, in constitutional typology, primary emphasis is on De Giovanni, and secondary emphasis is on Kretchmer, but Sheldon and Naccarati are not mentioned.

Despite the obvious (from American standards) over-emphasis on Nationalistic productions, this is a very fine, sound, and well-rounded volume. The book is appended by a series of examination questions covering each of the chapters, and contains 34 illustrations.

San Quentin. RAYMOND J. CORSINI.

ZIPF, G. K. *Human behavior and the principle of least effort. An introduction to human ecology.* Cambridge, Mass.: Addison Wesley Press, 1949. Pp. xi.+573.

This book is the work of a philologist who has surveyed many fields of knowledge in an effort to discover a general principle around which to organize the facts of human behavior. He believes he has found such a principle in the notion of Least Effort. But the reader, faced with the complications of terminology and the multiplication of concepts, seeks in vain for a satisfactory way to minimize the effort of comprehension.

Zipf criticizes briefly the various efforts of psychologists working with the principle of least work or least effort, and goes on to define the concept for his own purposes as the attempt to "minimize the average rate of work-expenditure over time." The principle is generalized to the fullest possible extent. It is said to encompass the following areas, among others: the economy of words from the point of view of the speaker, the economy of speech as sensory impression upon the auditor, the problems of "mentation," the problem of culture and of cultural language, social communication, the distribution of biological species, the explanation of biological evolution, the dynamic basis of the Freudian death wish and life wish, the economy of geography, and problems of economic power and social status.

It is beyond the limits of a review to attempt an evaluation of all aspects of this ambitious program. It is diverting, however, to try to take our bearings by examining the manner in which Zipf deals with certain problems which are familiar to the psychologist and which are presumably useful in studying behavior. What of motivation? Inasmuch as "the entire behavior of an individual is at all times motivated by the urge to minimize effort," treatment of the problem of motivation is greatly telescoped and the question of conflicting motivations simply vanishes. However, despite the overwhelming importance attributed to least effort as the dominating goal of all behavior, the author cannot completely escape the existence of other kinds of motivations. He does not recognize them as such, but speaks instead of the fact that much of a person's total behavior can be expressed in the form of the simple problem of *tools* and *jobs*. There is an attempt to match the tools to the job and to some extent the job to be done will be determined by the tools at hand. So the objective of winning a war will require that available peace-time tools be adapted to making tanks. On the other hand, "each human being is a set of sexual tools for which a reciprocal set of sexual jobs must be found." Thus sex drive would appear to arise from the economy-bent objective of using sexual equipment which would otherwise be idle, and if the logic of the situation is pursued, sex activity becomes an expression of thrift. This seems to be a kind of inversion of Freud.

In the attempt to figure out ways to minimize the possible average rate of work, intellectual and semantic processes are involved at all levels. Conflict, when mentioned briefly, is identified as conflict "in semantic systems and in problems of classification." A section headed, "Mental disturbances: the Need for Logical Self Consistency," discusses mental difficulties as involving "classificatory disturbance." Even physical traumata are laid at the door of semantics. Zipf says, "Whether in healing the wound itself or in going to a doctor for medication, the semantic system is trying to restore a logical self consistency." Turning to a discussion of the ego, we find this concept defined as, "A point in time-space which is to serve as a form of reference. This point is to be construed in the mathematical sense of a point, though it may be called an ego, or identity point, or a soul." This would seem to be an appropriate place to end the search for concepts that make some connection with the "identity point" of the average psychological reader.

This system of concepts is evidently not seriously anchored in observation of human beings as the psychologist knows them, though strangely enough there is in the book a great amount of empirical material which involves human behavior. There is, for example, an extended body of statistical data, showing distribution curves of a large number of phenomena, such as frequency of each word in James Joyce's *Ulysses*, the number of articles of various lengths in the eleventh edition of the Encyclopedia Britannica and in various newspapers, the frequency of trips of given distances made by trucks and cars, the frequency of marriages in which mates lived different distances apart, and so on.

Now these are acceptable data for social science and to question their importance on the basis that they are not those with which the psychologist most commonly deals would properly be called professional ethnocentrism. There can be value in establishing relationships between environmental factors and almost any aspect of behavior. Opinions differ about the necessity of writing into the equations of these relationships processes in the individual such as personality factors, drives, conflict, and the ego. It is scientifically acceptable to leave these factors out if accurate predictions can be made without them, but it is not acceptable to introduce them, as Zipf does, without making any contact with the large body of empirical psychological fact about personality and other aspects of the individual with which psychology deals.

HELEN PEAK.

Connecticut College.

RYAN, T. H. *Work and effort: The psychology of production.* New York: Ronald Press Co., 1947. Pp. xii+323.

Summary of this Review: I found the book to be both interesting and important. I shall comment chiefly on some of those assertions in it from which I dissent.

Consider Ryan's definition of Effort, *e*. (All the symbolic notation is mine.) Effort (e) is "the relationship between rate of performance" ($=P$) "and the capacity" (P_c) "of the individual at a given time."

Ryan strove for operational definition. Did he achieve it? He did if he indicated a set of operations that are (1) clearly conceived, (2) unequivocally denoted, and (3) capable of being performed. Otherwise he failed. Let us see.

Performance ($=P$)—in what task? Evaluated by what standard? In terms of useful Work? Power? Force? Action? . . . ? Is it to be expressed in terms of energetics, or perhaps rather in terms of some social concepts, such as how many letters are correctly typed in a minute, how many cards are correctly sorted in an hour, how many problems in mental arithmetic are correctly solved per calorie of heat output? Or, perhaps, in some terms not mentioned in this list? Are these rate-concepts interchangeable? If they are not interchangeable then each should bear a special name, not a generalized name used ambiguously. I say it regretfully, but it seems to me that the author carries this term through leaving it operationally undefined.

One may now ask when is the "given time" that Ryan puts into his definition? Are P' the actual performance and P_c' the maximum performance to be determined on the same date? If so, how does one distinguish between P' and P_c'? It seems to me that all we can observe on a "given" date is what the subject's performance P' then is. We have no means of finding out what it, at most, *could* have been. In other words we cannot determine P_c' his maximally possible performance, "or performance-capacity" on that date. Then, what meaning shall we assign to P_c'? Shall it imply the maximal performance-rate that the subject has *ever* yielded? Or perhaps the *average* of his highest performance-scores between certain ranges of dates? Or if not, what then?

Whatever arbitrary dating-system and spacing system we adopt, we get a score P_c'' that is logically capable of being exceeded at any time. Now, suppose that $P' > P_c''$. Operationally, it means only that the performance of the subject on one date surpassed his actual performance on some *other* specified date. (I say "other" because, as we have noted, his performance on one date *cannot* exceed his performance on the same date.) But now suppose we label P_c'' his "capacity" without using any qualifying adjectives or subscripts. Then we have said that his dated performance exceeded his undated capacity for performance. Since $P/P_c \leqq 1.0$ (wherein $1 =$ the index associated with maximal performance) this implies that P'/P_c'' can exceed unity, which contradicts the definition. So, we must either make some very awkward assertions, or else repudiate the definition. For it implies that a subject *can* make less than no effort at all.

This fuzziness of definition pervades many parts of the work. In my opinion it does not serve to make it easier to read understandingly.

In his first chapter Ryan presents a table (1, p. 16) that purports to show, for varying population-numbers, the "range of correlations which may be obtained by chance" if "no correlation exists in the population from which the group" (i.e., the sub-sample) "is drawn." Now suppose a correlation that falls outside this range. Should we then declare that it could not have been obtained by chance? Would it not be better to say that the probability of getting so large a value by chance is 10^{-n} in which n is the number in the probability series corresponding to the variance, or to the so-called critical ratio that we obtained? It is my considered opinion that questions such as these cannot be economically discussed with students who lack one year of collegiate mathematics and a year in sound statistical training. But on the other hand, I have seen no *useful* book on industrial psychology or generally in applied psychology that presupposed less training than that.

I differ from the author in his presupposing that decrement of performance measures, in general, the effectiveness of an agent of impairment, such as prolonged labor, deprivation of sleep, deprivation of oxygen, or dosage of alcohol or a drug. For the presupposition denies by implication—but yet effectively—that the living organisms is a compensating mechanism, though it really is. Like other compensating mechanisms it tends toward homeostasis, either at one "level" or else definitely at some other level. In studying impairment one might, of course, measure decrement of performance, but one will "miss the boat" if he overlooks observing and taking into account the various devices which the organism uses to maintain, in the face of progressive impairment, something like an equilibrium-state or a set of equilibrium states. Therefore I have to regard part of Dr. Ryan's work as being maldirected.

My space being limited, I shall now mention a few matters as briefly as I can. On p. 269 the author presents an "experience-table" incorrectly referred—though doubtless accidentally—to an article by Dodge and Romig. It shows three hypothetical distributions of accidents per operator. The first is based on a rate $a_1=2$ accidents per operator; the second on a rate $a_2=0.5$. Both these distributions are Poissonian. The third distribution composed of the sums of the other two, is non-Poissonian as it necessarily is if its components, each being Poissonian, are based on different accident-rates and *therefore* on different variances.

One would like to reverse the process—I once tried it for a short time—and analyze a given non-Poissonian distribution into two or more Poissonian components. It can be done—either by means that correspond to the analysis of a Fourier series, which will yield an exact result; or by the method of moments, as the late Percy W. Cobb proved in an article that still lies unpublished. But in either event, one may expect to get (1) some negative fractional populations, and (2) some negative exponents, either real or imaginary, to satisfy the formula. The problem is like that of the "vector-analyzers" or the "factor-analyzers" in psychology: one can find a unique *resultant* appropriate to the problem described if one is given the various components; but one cannot find *any unique* set of components from a given resultant. And, one is left with a group of operators who as a group are definitely abnormal, only the rest of the total distribution will contain some that belong in the same accident-classes as these. The method does not show which are which.

I must now record exceptions to what the author says about some work which I personally directed. Part of it was published jointly, part severally, but I can speak authoritatively on both parts. It is true that the Connecticut sample population which Cobb compared with a Poissonian distribution to find the correlation with accident rate that "the best of all possible" tests would yield was not a "random sample."

Many of the subjects had been ordered to take the Lauer and DeSilva tests because they had been recently involved in accidents and the question of their re-licensure had been raised. It is not true that this kink in the sampling was in any way objectionable—unless one insisted, against Cobb's warning, on treating the sample as if it were different from what it was, and was plainly declared to be. No wise person commits extrapolation in a public place.

It is true that the distributions which Johnson and Cobb used to demonstrate a regression toward mediocrity in accident-rates of the same operators in two different periods were not normal; it is also true that in computing the probability of the differences between two proportions we used a formula that presupposes that such differences are normally distributed about a mean of zero. But the largest chance-probability that we found by the conventional formula was of the order of 10^{-41} and I wonder what a correction of a second order would be worth? Also, as Ryan says on p. 292, I did not take notice of Dunbar's statistical work on psychosomatic problems in accident-making. This happened partly because most of my contribution to this work was done up and written before Dunbar's work was published. I make bold to say now, however, that I do not set a high value on her statistical procedure.

I repeat, this book is interesting and valuable. But to get the maximum value from it one must read it critically rather than with a view to gulping and regurgitation.

H. M. Johnson.

Tulane University.

Buros, O. K. *The third mental measurements yearbook.* New Brunswick: Rutgers Univ. Press, 1949. Pp. xiv+1047. $12.50.

Those who have been working for some time in the field of measurements are already familiar with the plan and purposes of *The Mental Measurements Yearbook.* Newcomers in this field will soon get the information for themselves, because the *Yearbook* has established itself as one of the necessary volumes in personal and departmental libraries. As in earlier volumes, there are "frankly critical" original reviews of tests and of books on testing, as well as excerpts from reviews which have been published in other places, the aims of the volume being not only the presentation of information about tests and testing, but the encouragement—indeed, the compulsion—of test makers and writers to do better work.

An attempt has been made to list all commercially available tests and books for testing, which have been published between October 1940 and December 1947. In addition, a few which were published at earlier dates and which were missed in earlier editions are also reviewed. Six hundred sixty-three tests and five hundred forty-nine books are referred to. The whole series of *Mental Measurements Yearbooks* constitutes the most complete bibliography available, and so far as this writer knows, the only extensive one which gives critical evaluations. The term, "Mental Measurements," it should be noted, includes the areas of tests of educational achievement, vocations, and the like.

Those who contributed to the *Yearbook* were charged to be "frankly critical with both strengths and weaknesses pointed out in a judicious manner." While the invitation emphasized strengths as much as weaknesses, the prevalent connotation of the word, "critical" has led many of the reviewers to give the major portion of their attention to shortcomings. Perhaps this is desirable in order to attain the objective of securing better performance in the making of tests.

It is not to be expected, of course, that the products of three hundred and twenty reviewers should be of equal quality. Some reviewers are more competent than others and some take their assignment more seriously than do others; but, in general, the work has evidently been done conscientiously.

If one were to take the amount of space devoted to give tests or books as evidence of their relative importance, it might lead to erroneous conclusions. Much discussion may mean importance; or it may mean great controversy, or unestablished techniques, or current interest. Rorschach testing and projective techniques generally command many pages of the *Yearbook*. Stoddard's *Meaning of Intelligence*, judging from the nature of the reviews, is given much more attention than it deserves; and the editor might have been wiser, as well as more modest, if he had reduced the number and length of excerpts of reviews of *The 1940 Yearbook*. Here, as in many other instances, the duplication of matter saying the same thing does not add to the value.

The *Yearbook* is carefully indexed in several ways which enhance its merit, both in referring to its own subject matter and as sources of information to test users. These indices are: a periodical directory and index; a publisher's directory and index; an index of titles; an index of names; and a classified index of tests.

University of Nebraska. D. A. WORCESTER.

SUPER, DONALD E. *Appraising vocational fitness by means of psychological tests*. New York: Harper, 1949. Pp. 727.

Dr. Super has contributed a book important to psychologists in general clinical practice, particularly those in the areas of personnel psychology. The author has attempted to answer the question, "what does this test, and the score made on it by this person, tell me about his vocational promise?" Actually the publication goes far beyond the answers to such questions. The reviewer is especially impressed with Dr. Super's willingness to venture clinical judgments based on his long clinical experience. The author has also understated his scope by using the term *vocational* in his title. The Columbia group has tended to use this term in a very broad sense, and it is thus that Super has used it. For those prospective readers who, like the reviewer, consider *vocational* an extremely limited concept, Super might will have substituted *psychological fitness*.

Appraising Vocational Fitness should be a widely used book. It is an essential addition to the libraries of clinical psychologists whether they be engaged in industrial or educational applications or in the practice of psychotherapy. The administrator who employs or works with the clinical psychologist will find pertinent information. Because of the somewhat informal methods of presenting materials, upper division as well as graduate students should find it readable and interesting as a text or reference. The reviewer does not believe that Super has attempted to replace such books as Bingham's *Aptitudes and Aptitude Testing* or has introduced a volume in competition to such a publication as Buros' *Mental Measurements Yearbook*. Rather, he has contributed interpretive data which make other presentations in the field of psychological measurement more useful.

The book moves at a fairly rapid pace; its evaluation of counseling theory and practice appears sound. Super's comments on Rogerianism and other approaches to counseling are sensible and objective. The number of years during which the manuscript must have been in preparation is evidenced by the carefulness of the workmanship and the comprehensiveness of the bibliography. Significant research has been carefully selected and well presented. Especial attention is called to chapters XVI, "The nature of interests," XXI, "Using test results in counseling," and XXII, "Preparing reports of test results." These chapters contain pertinent materials which, in the opinion of the reviewer, are presented well above the general level of many such presentations. Readers will find Appendix A, "Statistical concepts," a valuable review of oft forgotten basics.

Some readers will hope with the reviewer that when Super comes to the inevitable revision period he will have at hand a solution to the problems of consistent and psychologically acceptable terminology. Some of the chapter heads follow the general trait names so much used by clinical counselors with their clients. Examples are "clerical aptitudes," "musical talents," and "mechanical aptitude." However, a separate chapter is presented on "Manual Dexterities" and another on "Spatial Visualization," concepts which are frequently presented in the literature as factors or traits which are aspects of mechanical aptitude and ability. This is not criticism but rather a hope that *Appraising Vocational Fitness* when revised may help us to escape from a semantic wilderness.

In conclusion, this book is highly recommended as a professionally sound, well-written, authoritative and thoughtful presentation of the use of tests in clinical psychology.

MILTON HAHN.

University of California, Los Angeles.

MATHEWSON, ROBERT H., *Guidance policy and practice.* New York: Harper, 1949. Pp. xiv+294.

It is probably generally agreed that sound attempts at formulation of a framework of fundamental theory for guidance are highly desirable. This book represents such an attempt. Basic assumptions and concepts are dealt with extensively in discussing guidance policy and practice. The approach is broad—perhaps too much so.

Discussion is generally on a high abstract level. It is true that many of the topics covered have little or no related research data on which to base discussion. The reviewer feels, however, that this book is definitely lacking in presentation of supporting evidence where such is available. The typical school administrator may have difficulty in following the discussion while the guidance specialist will likely be disappointed in the lack of scientific supporting evidence.

Considerable discussion in this book deals with the frontiers of thinking in guidance. It is easy to criticize unduly any person making such an attempt, but even if only a portion of the theoretical framework as presented ultimately proves valid, the attempt should be considered worthwhile.

The book is divided into four parts. Part I covers fundamental factors such as basic concepts, the needs of individuals, the needs of society and cost of guidance programs. Part II includes the nature of guidance policy and its implementation in the school and community. Part III presents the issues and problems in guidance, such as the responsibilities of educational systems and society, types of personnel and training required, and the role of the counselor. Part IV discusses the future of guidance and the need for cooperation of education, as the primary medium, with the total community. The following conclusion is presented: "Results of our study of the needs for guidance in American Schools and community, and of the forms of guidance service required to meet these needs, all point toward a single imperative requirement: a national policy and strategy of guidance and personnel service in education" (p. 265).

Mathewson did not sidestep difficult philosophical and psychological issues in writing this book, but that does not imply that all such issues are treated to the satisfaction of all readers. The book, however, is not intended to be a final blueprint of a fundamental theory for guidance. For the person who is interested in the formulation of guidance policy and practice, this book should prove to be a thought-provoking review of many of the problems which have not yet been solved.

FRANK M. FLETCHER, JR.

The Ohio State University.

DONAHUE, WILMA T., COOMBS, CLYDE H., & TRAVERS, ROBERT M. W. *The measurement of student adjustment and achievement.* Ann Arbor: Univ. of Michigan Press, 1949. Pp. 256. $3.00.

The Institute for Human Adjustment at the University of Michigan organized a conference in June, 1947, on the topic of this title. The conference was intended to be of interest to psychologists and educators and to students in these fields. In contrast with many educational conferences on the subject of guidance, the lecturers apparently were asked to assume reasonable training on the part of all members of the audience so that the content of the lectures would be at a much higher level than is common. This volume presents these lectures.

Gilbert Wrenn's opening lecture reviews the status of the guidance movement, the development of personnel work, and the stabilizing of the profession.

The next four lectures are on the general topic of measurement of adjustment. Percival Symonds and Murray Sherman review a personality survey which they conducted in a junior high school, from the point of methodology and its contribution to good counseling. Harold Edgerton brings the reader up to date on some of the technical aspects of measuring instruments in guidance. Max Hutt speaks on projective techniques in guidance and Wilma Donahue closes the section with a discussion of the training of guidance workers, especially with reference to psychodynamics of personality.

The next group of lectures comprises reviews in the field of achievement testing. Irving Lorge reports on trends in the measurement of achievement. Warren Findley emphasizes the development of large-scale achievement testing, describing some of the organizational and technical problems of programs like the College Entrance Examination Board, the Iowa tests, and some of those used in the military services. Harry Rivlin presents an analysis of the teacher's role in achievement testing which is a proper emphasis, since in these days of the expert test-maker and the centralized testing service, the teacher sometimes is forgotten. Arthur Traxler's presentation of diagnostic testing in relation to remedial teaching points out the important idea that diagnosis is an analytical process properly carried on by competent people using a wide variety of sources of information. Tests provide much of that information, and the more specific these tests become, presumably the more unique contributions each can make to the final diagnosis.

Another group of lectures is concerned with the topic of prediction of success. One of the more comprehensive is that by Robert Travers who reviews research on the prediction of academic success. This chapter has a bibliography of 272 items which ought to be of interest to persons working in this field. When one thinks of Walter Dearborn one expects a lecture on the psychology of reading. This time Dr. Dearborn delves back into his rich experience and comes up with some notes on the relation of a student's background to his school success. In the anecdotal manner, but quite pointedly, Dearborn extracts several ideas which should make those that think they know all the answers a little less smug. David Trout reviews attempts to correlate academic achievement and subsequent success in life. He reviews the research and problems of research involved in using such criteria as eminence, wealth, occupational status, and so on. One of the most important points he makes is that colleges and schools might do well to make researches as to why persons who fail in their institutions somehow succeed in life. George K. Bennett's chapter outlines some of the major problems in guidance with respect to vocational and technical training. One of his main points is that guidance must be early enough so that practical decisions can be made. The choice points are probably early in the junior high school and again at about the eleventh grade.

The major address in the last section is that by Clyde Coombs, who emphasizes a new slant in the measurement of psychological traits and outlines a program of research to investigate his ideas. At the present time traits are measured and a person is assigned a status score. Two persons can have the same status score on the basis of passing different sets of items in the test used. Coombs suggests also that a dispersion score be developed as a meaningful concept to help describe persons. He also points out that we can develop a trait status score and a trait dispersion score, these scores providing for us a better understanding of the traits themselves. In group studies the dispersion of the status scores and the dispersion of the individual dispersion scores become units of importance.

Jacob Orleans was asked to close the conference with a description of the guidance program of the future, which chapter naturally becomes a statement of what the author thinks will be good standards of professional practice in the field of guidance.

The writer of this review heard all of these lectures and talked with both trained psychologists and educators and with graduate students and teachers in attendance. His impression was that such persons appreciate a conference in which the level of presentation is kept high enough to be stimulating to all and which gives one a feeling of having moved forward in one's own development. Naturally, not all of these chapters are major contributions, but there is enough in this book to warrant careful reading by specialists and practitioners in the field of measurement.

HAROLD SEASHORE.

The Psychological Corporation.

HURLOCK, ELIZABETH B. *Adolescent development.* New York: McGraw-Hill, 1949. Pp. viii+566. $4.50.

This book summarizes the principal studies on adolescence reported up to 1948. It is organized around the usual topics of adolescent development, behavior, interests, attitudes, and personality, in a conventional but, at times, obscure style of scientific reporting. Because of the large number of data reviewed in this book, the reader must remain vigilant lest he miss the forest by too much preoccupation with the trees. Topic headings are judiciously and generously used throughout the book. Unfortunately, no chapter summaries are included, which omission makes for needless confusion and labor on the part of the reader. Unfortunately, the basic principles are not set off from their supportive argument, making it necessary for the reader to ferret them out for himself. These mechanical defects, along with an absence of critical discussion of methods used in the study of adolescence, will provide those instructors who use this book as a text with a good deal of work.

The outstanding feature of the book is its insistence that adolescence is an age of transition, rather than an age of transmutation. Contrary to popular opinion, adolescence is not an age of miracles. The basic behavior patterns of childhood do not suddenly become transmuted by some fortuitous alchemy into fully mature, socially approved reactions. Throughout this book, evidence is presented "to show that traits present in childhood will become more deep rooted with the passage of time and that undesirable traits, instead of disappearing, will be even more undesirable." This is a lesson parents, teachers, ministers, child welfare workers, and all others responsible for youth should learn and learn well. It is a lesson that is persistently but, often-times, obscurely taught in this book.

The author advances the idea that most of the so-called abnormal behavior of adolescents is not really abnormal at all when judged by the criteria of the adolescents' peer culture. As the author points out, most of these so-called abnormal phenomena ". . . . are an accompaniment of growing up and, as such, should be regarded as normal and natural." Unfortunately, the reader must worm his way through a mass of conflicting data to arrive at this notion. No clear-cut studies are presented which expressly support the author's contention, though her point of view is ably supported in the incidental findings of a number of different studies.

The reviewer feels that some important studies have been neglected in this book. For example, the experimental studies on level of aspiration are overlooked, as are the field studies on social climate of Lewin's students, and the frustration and conflict studies of Rosenzweig, the Yale group, etc. In fact, in the reviewer's opinion, the principal shortcoming of the book is its excessive use of so-called "critical-ratio" research. More elaborate experimental studies than those of the usual questionnaire and attitude scale type should have been included. While not expressly aimed at the study of adolescence, there is much valuable material in these more elaborate studies which throws light on the adolescent.

There is, throughout the book, a sort of native faith in simple cause-effect behavioral relationships. For example, on pages 115–116, the author states that "The most common causes of adolescent suicides are (1) disappointment in love, (2) failure in school work or in jobs, (3) maladjustment to work, (4) homesickness, (5) maladjustment to parents, and (6) intellectual doubt and anxiety." While these factors may be present in adolescent suicides, they alone surely do not cause it. Our science has demonstrated repeatedly that behavior is over determined in the sense that there are many interacting causal factors in all aspects of behavior. This book overlooks basic behavior dynamics all too often.

There are some inaccuracies in the book. An example is on page 486 where the author states that "Insanity in the forms of dementia praecox, hysteria, and manic-depressive psychosis is becoming increasingly frequent." Apart from the fact that it is incorrectly regarded as a form of insanity, hysteria is definitely on the decrease. Whether the author should be so dogmatic about the increase of dementia praecox and manic-depressive psychosis in view of the excellent study of Landis and Page is to be debated. In the reviewer's opinion, a reported increase in psychosis should be accounted for in terms of the limitations of the data which tend to show the increase.

The reader is taken back quite a few years in the treatment of the physiological concomitants of behavior. One gets the impression that this concomitance is a one-way affair with disturbed physiological functions affecting behavior, but not vice versa. This is regrettable in virtue of the more recent psychosomatic emphasis.

On the whole, the book is well done and the reviewer feels it would make a satisfactory basic text for a course in the psychology of adolescence for students who have had a preliminary course in general psychology. It should also prove a valuable reference for other courses like child psychology, educational psychology, and social psychology. The busy specialist will appreciate its scientific flavor and will find it a convenient way to review the current literature (up to 1948) on adolescence. In the reviewer's opinion, this is not a book that could be recommended to parents, due to its scientific style of reporting. One final point in favor of the book deserves mention. The author indulges in practically no moralizing. This is especially commendable in a field where most popular books are disgustingly moralistic.

ROBERT P. FISCHER.

Marietta College.

ESCALONA, S. K. *An application of the level of aspiration experiment to the study of personality*. Teach. Coll. Contr. Educ., No. 937. New York: Bureau of Publications, Teachers Coll., Columbia Univ., 1948. Pp. viii+132. $2.10.

This monograph is of limited interest. The differences in behavior in the level of aspiration situation between two groups of school pupils are described—the control group selected as well adjusted in school and the experimental group selected on the basis of maladjustment in school.

The study differs from previous experiments on level of aspiration primarily in the conceptual system in which the results were analyzed. The author points out that no significant relations have been discovered between quantitative L. A. scores, especially discrepancy scores, and certain personality variables. It is her contention that relations can be discovered if the experimental observations are interpreted within the theoretical constructs of vector and topological psychology. Thus we find that 30 pages are devoted to quantitative results which lead to seven conclusions as contrasted with 52 pages of qualitative results leading to ten conclusions.

The work can be adversely criticized on the ground of the selection of subjects. With such small samples, 19 in each group, the failure to equate for the parameters which might possibly produce differential responses between the two groups becomes of major importance. For example, 90 per cent of the control group and only 26 per cent of the experimental group were girls. And again, in attempting to explain why the aspiration level of the maladjusted group was initially higher than that of the control group, she states (p. 31) that the majority of the control group were drawn from one particular school, "where stress is laid on orderliness and a systematic approach to a new task."

The composition of the monograph follows the accepted pattern of presenting an experiment. Nine chapters follow in this order: statement of the problem, methods and procedure, results and conclusions. There is a bibliography of 49 titles, 40 of which are concerned with other level of aspiration studies.

RICHARD T. SOLLENBERGER.

Mount Holyoke College.

LIEF, ALFRED, *The commonsense psychiatry of Dr. Adolf Meyer*. New York: McGraw-Hill, 1948. Pp. xvii+677. $6.50.

This book consists primarily of 52 of the many papers written by Adolf Meyer during the approximately 50 years of his work as a psychiatrist in the United States. Most of the articles are reprinted, sometimes in abridged form, from well-known books and journals, but many derive from relatively obscure sources, such as privately printed reports and addresses.

The 52 articles are classified in twelve sections, which have such titles as "Preparation in Europe," "Action in Worcester," "Training the Doctors," "Science of Man," etc. Lief's contributions to the book, aside from his selection of the articles, include a foreword, an introduction to each section, reference notes, a glossary of medical and philosophical terms, and an index.

The aim of the book is "to give the average person a better practical understanding" of Meyer's viewpoints, "to portray the evolution of a psychiatrist and of his thinking and work." It is, Lief says, "an exposition of American psychiatry and at the same time the picture of a physician in action" (pp. xi–xii). To this end, the arrangement of the twelve sections is basically chronological. Each section, as noted above, is preceded by an introduction which is both descriptive and interpretative, and which serves to put in proper perspective the group of articles which follow. These introductions, taken together, provide in themselves an excellent account of the development of Meyer's "psychobiology," and of his work as a teacher and a leading figure in the psychiatric world.

In the opinion of the reviewer, the book is not without at least minor defects. For one thing, the fact that it attempts to combine a chronological narrative of Meyer's professional activities with a presentation of his various concepts and theories deprives it of any single organizing principle. Also, the reader will not find, even in Lief's introductory discussions, any clearer exposition of Meyer's psychobiology than is already contained in such familiar articles as "The Psychobiological Point of View" (in M. Bentley and E. V. Cowdry, *The problem of mental disorder*. New York: McGraw-Hill, 1934; reprinted in the present volume). Furthermore, the various articles in the book vary considerably in their interest to most readers. Some, especially the earlier papers, are of value chiefly because of the light they cast on the successive stages of Meyer's intellectual and professional growth. Others serve mainly to reveal the huge scope of his interests and activities, *e.g.*, in medico-legal problems, the mental hygiene movement, the teaching of psychiatry, the preparation of case histories, birth control, etc. It might be added that the attitude of Lief in this effort to present Meyer as psychiatrist in his entirety is that of an admiring pupil rather than of a critic.

Nevertheless, the book is indispensable to anyone who wishes to obtain a thorough knowledge of Meyer's thought and work, and of the extent of his influence on American psychiatry, without himself undertaking the task of locating and bringing together scores of widely scattered papers. Meyer himself was either too occupied with his work as physician, teacher and executive, or too disdainful of what he called "dogma," ever to undertake a reasonably complete and authoritative version of his theories and his psychiatric methods. In the absence of such a version Lief's work is very welcome indeed to the interested student.

As the book makes clear, Meyer's influence on psychiatry in the United States undoubtedly has been very great. It is the more surprising—though Lief does not note this point—to observe how small his direct influence on psychology seems to have been. Even in the textbooks of abnormal psychology, psychobiology either is only briefly described or, more usually, is not mentioned at all. Also, the references to Adolf Meyer in the indices are not numerous. In eight fairly recent texts, the number of such references ranges from two to twelve, with a median of 3.5, and many of them are to Meyer as author only.

It is possible that psychobiology has had, indirectly and by a sort of infiltrative process, some share of the influence on psychology which its affinity with much respectable psychological theory would seem to merit. For example, it may have played some role in furthering the acceptance within psychology of such conceptions as those of the organism-as-a-whole, the value of the personal history approach in the study of an individual, the importance of social factors in shaping personality, the emphasis on functional and environmental factors rather than on organic and genetic ones for the explanation of behavior, and so on. However, if psychobiology has had such influence, the fact is not explicitly recognized in most psychological writings.

Perhaps one reason for this lack of direct, recognizable influence is the frequent obscurity and difficulty of Meyer's style, together with his employment of various ill-sounding "neologisms" such as ergasia, dysergasia, anergasia, and the like. Another reason may be the fact, already referred to, that Meyer himself never presented his whole theory in organized and systematic fashion. However, it is possible that the principal reason for Meyer's apparent lack of influence on psychology is that his concepts are too much like those with which most psychologists are already equipped. His "commonsense" psychobiology resembles in many important ways the more or less eclectic, and only loosely systematized explanations of human behavior to which a majority of psychologists adhere. In consequence, Meyer's theories—whatever novelty they may have had for psychiatrists—provide most psychologists with few if any new insights, and merely reaffirm concepts with which they have long been familiar. In the scientific world, as elsewhere, the most credit as well as the most censure tends to accrue to the dogmatic revolutionaries, among whom psychologists will include Freud and Adler but not, we think, Dr. Adolf Meyer.

LELAND W. CRAFTS.

New York University.

KANNER, LEO. *Child psychiatry.* (2nd Ed.) Springfield, Ill.: Charles C Thomas, 1948. Pp. xxiv+752. $8.50.

After passing through five printings this pioneer contribution to psychiatry and pediatrics appears in a completely rewritten edition. Its present form deals respectively with parts on history, basic orientation (age, physical conditions, environmental and psychological variables such as intelligence and attitude), clinical considerations (symptoms, diagnosis, therapy and prevention), and phenomenology. This final part includes a section on (1) personality problems arising from physical illness, (2) psychosomatic problems, and (3) problems of behavior. The largest section, problems of behavior, covers a wide range of syndromes such as complaints connected with vegetative functions, bodily manipulations, speech, sex, scholarship, delinquency, anxiety and neurotic symptoms, schizophrenia and suicide.

Treatment of such a wide scope of subject matter in several overlapping disciplines perforce offers difficulties to any scholar. If the writer is also a busy clinician the task is almost insurmountable. In many of these areas the existing bibliography is very extensive and drastic condensation and selection is necessary. Dr. Kanner exhibits a broad approach to most of the problems treated, quoting studies and viewpoints of neurologist, sociologist, anthropologist, and psychologist. Furthermore, his selection of studies and evaluation and integration of conflictory viewpoints in the many fields treated is a real contribution. Coverage of journal literature, however, indicates that the task is beyond the capacity of a single writer who represents one profession. This problem adds validity to the recent trend of several specialists cooperating in the publication of advanced psychology textbooks. The latest date, for example, in the bibliography on emotion proper is 1923. Half of the references on intelligence date from before 1926 and this section falls short in comparison with a standard undergraduate text in Child Psychology

The part entitled "Basic Orientation," which touches on psychological theory, is in marked contrast with the rich clinical chapters in the latter parts of the book. It is loosely organized and does not present a systematically consistent approach to human behavior. For example, both the views of Allport and Guthrie are espoused without indication of their differences. The disparity between these two parts of the book argues for continued and closer cooperation of the two professions. The psychologist working with theory of personality and experimental design of relevant laboratory studies, and the psychiatric clinician both can benefit by such a union. Moreover their students will find this mutual complementation imperative to adequate training.

Kanner is as eclectic as his Meyerian background might suggest. The second edition, however, shows less of Meyer's obvious influence. A chapter on psychobiology as such has been dropped and the emphasis on Meyer's unique terminology has abated. The volume will strike the contemporary clinical psychologist as conspicuously brief in the treatment of adjustive mechanisms and deeper psycho-dynamics.

In psychotherapy the emphasis is on release through projective and play techniques and by the establishment of a good relationship with therapist and those who influence the child daily. Stress is placed upon understanding the attitude of those who influence the child, and assisting them and the child to achieve better perspective and relationships. Therapy is neither highly directive nor non-directive. Therapeutic suggestions are included in the discussion of almost all syndromes, but implied in these discussions is the importance of all relationships within the child-environment pattern.

This book will continue to serve as a handbook for the clinician and the student interested in problems of children but beside it should stand a good text in child psychology and a book on psychodynamics.

FRED MCKINNEY.

University of Missouri.

DUNBAR, FLANDERS. *Synopsis of psychosomatic diagnosis and treatment.* St. Louis: C. V. Mosby, 1948. Pp. 501.

ALEXANDER, FRANZ, & FRENCH, THOMAS M. *Studies in psychosomatic medicine.* New York: Ronald Press, 1948. Pp. xiii+568.

Other than in title and technique of collaboration, there is little in common in these two treatments of psychosomatic disorders. Broadly stated, Dunbar's "Synopsis" is a guidebook for the general practitioner who seeks information on the role of personality factors in the varied categories of physical disorder. The organization of the volume follows Osler's *Principles and practice of medicine.* Chapters contributed by specialists deal with diseases of the vasomotor equipment, of skin, eye, ear, nose, teeth, endocrines and of the gastro-intestinal, respiratory, circulatory, genito-urinary and metabolic systems. Separate, short chapters describe accident proneness, functional disorders and industrial medicine. Coverage is clinical, with scattered references to research material. Each syndrome is accompanied by description of psychological etiology and therapy.

The second volume under review is a series of reprints detailing the research and experience of members of the Staff of the Chicago Institute for Psychoanalysis. Unlike its companion volume which seeks its unity in a nosology of disease, the work by Alexander and French finds its coherence in the analytic orientation of its contributors. Prefaced by a number of articles on general principles of psychosomatic problems as seen from the vantage point of psychoanalysis, the remainder of the book deals with the gastro-intestinal, respiratory, cardio-vascular, endocrine, muscular and miscellaneous diseases. Emotional and personality factors appear frequently in the titles of various papers.

For the psychologist, some interesting queries are raised both by the contents and the assumptions of these two treatises. In both, lip service is paid to the necessity for enriching the understanding and therapy of physical disorder by information on the personality organization and psychology of the patient. However a count of references indicates almost complete neglect of publications from systematic (academic?) psychology. Of the 432 bibliographic citations in Dunbar's book, only six are the work of psychologists. Alexander and French are equally merciless in their indebtedness to psychology, four of 408 references cited emanating from psychologists. (Of these four, two deal with the James-Lange papers on emotional theory.)

However interesting, alarming or revealing these statistics, it is more important to examine the orientation of the specialists from the medical field toward the ancient problem of body-mind relationships. Both volumes regard their contribution as a partial solution along monist lines of the old controversy. Alexander, quoting from the opening editorial of the first issue of the *Journal of Psychosomatic Medicine*, notes that "Emphasis is put on the thesis that there is no logical distinction between 'mind and body,' mental and physical . . . . It takes for granted that the psychic and somatic phenomena take place in the same biological system and are probably two aspects of the same process, that psychological phenomena should be studied in their psychological causality with intrinsically psychological methods, and physiological phenomena in their physical causality with the methods of physics and chemistry" (p. 3). In illustration, emotionally caused elevation of blood pressure is admittedly effected by a physiological mechanism but the psychological etiology must be studied psychologically through introspection or by verbal communication. It follows, according to our authors, that psychogenesis is a physiological process of excitation in the nervous system which can be studied by psychological methods because the excitations are perceived subjectively as ideas, emotions, etc. Scientific methodology for the study of these subjective phenomena was invented by Freud. "Personality research or motivational psychology as a science begins with him (Freud)" (p. 20).

What has happened in this argument to the original contention that psychosomatic medicine is an integrative approach, a novel solution to the body-mind problem? Alexander states the answer in explicit terms: "We may hope that sometime in the future we shall also know the concomitant physiological processes, but even then it is an open question whether we shall dispense with our psychological understanding" (p. 28). Thus the reconciliation of the mental-physical dichotomy is but a programmatic foot-note with some doubt as to its ultimate validity. Psychosomatic medicine solves the controversy by elimination, for the present, of interest and activity in the somatic area and by sole emphasis on the analytic techniques for the study of so-called subjective factors in disease. Such semantic acrobatics make one hesitate to accept the claims of psychosomatic medicine in the field of systematic understanding of human behavior. For the psychologist it is revealing to note the disregard in the thinking of these specialists of the whole body of psychological research dealing with adjustment and development, the psychology of learning. Surely if one deals with personality processes, learning is a central factor. Its absence in the pages of these volumes reflects the poverty of training in general psychology which pervades psychiatric education. The book by Dunbar, although not as neglectful of psychological experimentation and research, still displays similar tendencies. One reads with pleasure references to Pavlov, to experimental neuroses in animals, to developmental patterns. Attention is called to the need for teamwork among psychiatric and psychological workers. Personality profiles are listed among the newer techniques available to psychiatrists but broad systematic treatment of human behavior is again lacking.

The foregoing criticism of psychosomatic thinking is based on the programmatic claims and operational concepts of the field. This does not imply that psychologists should neglect the body of research here described or fail to profit from the immense clinical experience of the psychosomaticists. Dunbar's volume has much of value for us, particularly in its coverage of the material in the medical literature. The first chapter, to cite but one instance, fills a gap in the picture of human developmental sequences. In this section, Hooker describes the chronology of behavioral events in the human embryo from six weeks of age to birth. Sontag, of the Fels Institute, does a thorough job on research in the area of predisposition to emotional and physical disorder. Some novel observations on the influence of habit systems on dentition are contained in the chapter by Dr. E. Schneider. This introduces a new and stimulating area of cooperation between psychologists and dentists. In general, the volume by Dunbar will particularly repay those who seek more extensive information on the present status of research on psychological factors in the disease process broadly conceived.

Positive evaluation of the book by Alexander and French is made difficult on the one hand by the unifying analytic orientation of the authors and by the diverse nature of the papers presented. Once granting the validity of the assumptions underlying the orientation, the research results are of consequence. However, if one feels hesitation, criticism is limited to the potentialities of the problem for those with divergent bias. Grinker's chapter, by way of contrast, contains a valuable and informative anatomical and functional treatment of the hypothalamus. The concept of encephalization of function is advanced as part of an integrative level approach to psychosomatic problems. For the most part, however, the chapters contain little more than case histories drawn from one to a half-dozen instances to illustrate typologies and etiologies previously hypothesized by the members of the Chicago Institute. Alexander, treating the problem of gastro-intestinal disorder, establishes three types on the basis of a reported population of 9, 5 and 5 cases. With such statistical attenuation, one looks with some caution on such statements as "In the great majority of our cases, we found an inner rejection of passive-receptive and oral-aggressive tendencies" (p. 112). Whatever value lies in these chapters is to be found mainly in the vast number of experimentally possible hypotheses as a base for future research.

BERNARD F. RIESS.

Hunter College of the City of New York.

Vol. 47, No. 3 May, 1950

Psychological Bulletin

PARTIAL REINFORCEMENT: A REVIEW AND CRITIQUE

WILLIAM O. JENKINS
Harvard University[1]
AND
JULIAN C. STANLEY, JR.
George Peabody College

I. Introduction

Little more than a decade has elapsed since the majority of psychologists interested in learning theory first became aware that there are important differences between continuous and non-continuous reinforcement, particularly when resistance to experimental extinction is used as a measure of response strength. During the 11-year period from 1939 through 1949 at least 30 experimental studies were reported in which the latter kind of reinforcement played a prominent role. Various animals, including humans, were reinforced on some but less than 100% of a series of trials or responses in situations where many kinds of behavior were measured. The very recency of these experiments and their diversity have resulted in a considerable body of uncoordinated data that needs to be analyzed critically and presented in systematic form to serve both as a current summary and a guide to further research.

Although "partial" is in certain respects not a thoroughly suitable adjective for describing the kind of reinforcement with which this paper deals, it has been widely used to indicate various combinations of frequency and pattern, parameters that can be manipulated independently. "Intermittent" and "discontinuous," two other words that sometimes take its place, both connote periodicity, while "partial" will accommodate the important random pattern with ease. "Continuous reinforcement" is used to cover the 100% case, "regular" implying periodic and most commonly employed. As discussed in this paper and the studies upon which it is based the term *"partial reinforcement'*, to extinction after partial as contrasted with continuous reinforcement must be included in this treatment.

---

[1] This research was facilitated by the Laboratory of Social Relations, Harvard University.

Partial reinforcement seems to be a basic condition of everyday behaving, for the rare instance is the presentation of reward each time a response occurs. (In this context reward is used to cover both the primary and secondary situations.) Examples, somewhat oversimplified, are easy to find: The animal trainer feeds his dogs only after several occurrences of the desired act, the clerk receives his check only once each month, the psychologist publishes his findings only after prolonged endeavor, the gambler continues to bet despite an infrequent payoff, *ad infinitum*. Continuous reinforcement seems to be the exception, not the rule.

Both practically and experimentally, partial reinforcement is a worthwhile procedure. It retards satiation effects and permits continued training over long periods and large numbers of trials. Accomplished animal trainers rarely feed their subjects after the completion of a single phase of an act, having hit on the efficiency of the method so that weeks of training can be condensed into a few days.

In brief, partial reinforcement is such a ubiquitous phenomenon that one wonders why psychologists for so long failed to examine its consequences.

The breakdown of the present paper is as follows. First the various techniques for presenting partial reinforcement are discussed. Included in this methodology section are comments and criticisms of partial reinforcement experiments, regarding problems in design and variable control. The four major sections of the paper deal with acquisition, post-learning performance, extinction and theoretical considerations. Some of the practical implications of partial reinforcement are indicated.

II. QUESTIONS OF DESIGN AND METHODOLOGY

There are certain methodological considerations, not conveniently discussed in connection with the later sections of this paper dealing with acquisition, performance, and extinction, that are of sufficient generality and importance to warrant separate treatment. A number of these are discussed in the following paragraphs.

Techniques of presenting partial reinforcement. The two major parameters in varying schedules of reinforcement are *frequency* and *pattern*. The distinction between them has sometimes been ignored and the two characteristics have been allowed to vary together, although the problem of patterning is present whenever frequency is less than 100% and greater than 0%. Mowrer and Jones (47) have pointed out that on several occasions experimenters have compared results from one group rewarded continuously and regularly and another rewarded discontinuously and irregularly.

With regard to frequency the whole gamut may be run from continuous reinforcement to reinforcing not at all (extinction). In practice, the working limits have been 75%–80% (**1, 3, 17, 66**) to 1% or less (**21, 34, 35, 36, 41, 62, 63**). Regardless of frequency, pattern may be regular or irregular.

Another way of considering methods of presenting partial reinforcement is to refer frequency and pattern to the continua on which they are varied in the experimental situation. These points of reference are *time* and *number of responses*. The reinforcing stimulus may be presented according to a regular or an irregular schedule in time at a high or low frequency. Similarly, frequency of reward may be high or low in terms of number of responses, with the pattern either regular or irregular. The most systematically studied case of regular reinforcement in time is Skinner's (**62**) periodic reinforcement.

The irregular time schedule, known as aperiodic reinforcement, has been recently developed (**63**) and employed (**14, 35, 36, 41**). By this method a rat in the bar-pressing apparatus might receive reinforcement for a response on the average of once every five minutes, with a range from one to nine minutes and a step unit of one minute. The intervals are presented in a randomized order. An asset of this procedure is that it precludes the decline in strength found with prolonged periodic reinforcement. The organism is given no basis for discriminating when reinforcements are, or are not, to be presented in time as in the periodic technique (**62**).

Presenting reinforcement for a regular number of responses has been called "reinforcement at a fixed ratio" by Skinner (**62**). In this situation the subject is given the reward only after it has accomplished a certain number of units of the behavior required, e.g., 50 bar-pressings. Skinner (**62**) has varied frequency from a reward for 16 responses to one in 192 bar-pressings.

The aperiodic counterpart in number of responses may be called random or variable ratio reinforcement. The organism might be rewarded on the average of once in 50 responses, with a range from one in 10 to one in 90 and a step interval of 10. This technique has been widely used with apparatus such as the runway and the T-maze (**1, 12, 13, 28, 29, 30, 32, 47, 59**).

A variable confounding the administration of partial reinforcement schedules is the type of experimental set-up employed. The situations that need to be distinguished are (1) those in which neither the experimenter nor the environment restricts the responding, e.g., bar-pressing. Throughout this paper we refer to this instance as "free responding." (2) A different situation is one where opportunities for responding are controlled by the experimenter, i.e., where trials are involved. The latter subsumes single-response instances such as the straight runway and alternative-response (T-unit) or multiple-response set-ups, e.g., various combinations of lights providing the stimulus patterns for several differential responses. The free-responding situation may be converted to the trial instance by withdrawal of the critical aspect of the environment (e.g., the bar) after each response (**52**).

One of the major differences between the free-responding and trial situations lies in the presentation of partial reinforcement according to a time schedule. The latter loses much of its meaning where the set-up is broken down into discrete trials. Where the time continuum is the basis for determining trials (viz., **50**, p. 41, and **2**), however, variations of reinforcement in time apply directly. Other differences lie in frequency and pattern. Typically, a randomized schedule of 50% reinforcement has been employed in the trial situation. Considerably lower frequencies administered regularly or irregularly have been used in the free-responding set-up. Methods of presenting partial reinforcement in the two situations are summarized in Table I.

TABLE I

Techniques of Presenting Partial Reinforcement

Variations of Reinforcement	Situation			
	Free-Responding	*Trials*		
		Single Response	Alternative Responses	Multiple Responses
Time				
Regular	Periodic reinforcement	Time variations not ordinarily used		
Irregular	Aperiodic reinforcement			
Number of Responses				
Regular	Fixed ratio reinforcement	Fixed ratio reinforcement		
Irregular	Random ratio reinforcement	Random ratio reinforcement		

Sequence of experimental conditions. One of the basic dilemmas facing the investigator of partial reinforcement is that of getting behavior started under a partial schedule. For example, it is clearly impractical to try to start rats from scratch on a 10-minute periodic schedule in the bar-pressing situation. Even with exceptionally high drive it is quite probable that because of the infrequency of reinforcement little or no conditioning will result. On the other hand, if continuous reinforcement is employed to condition the behavior, pure partial reinforcement in conditioning does not occur. Clearly, behavior must attain a minimal level of strength before anything can be done; continuous reinforcement is usually employed.

In a unique experiment Keller (38) has attacked this problem. After receiving enough continuous reinforcements to insure adequate strength of conditioning, one group of rats in the bar-pressing apparatus went through a sequence of continuous followed by periodic reinforcement. The other group had the conditions reversed. As measured by resistance to extinction there was a significant superiority of the group with continuous reinforcement just prior to extinction, but this difference held up for only the first five minutes of extinction. (Another recent study (9) supports this finding; see Section V.) Responding in the remaining 175 minutes of extinction was not significantly different for the two squads.

From these data it appears that the sequence of operations, typically continuous followed by partial reinforcement, may introduce a constant error if the extinction period is truncated. Otherwise, it seems unlikely that this variable will be critical in the free-responding situation.

In experimental instances involving trials, several experimenters (**3**, **7**, **12**, **13**) have been able to start their subjects off with a partial schedule. It should be noted that in these studies the frequency of reinforcement is usually high, of the order of 50%, and that ordinarily some form of preliminary training has been given.

In several studies such unusual sequences of experimental conditions have been employed that it seems advisable to discuss them as methodological points. One of these is Humphreys' (**29**) study of "expectations" in which he trained 78 students under 100% "reinforcement," extinguished the response, and then used the *same* subjects immediately as a 50% reward group with conditions otherwise unchanged.[2]

Brogden (**1**) reinforced three groups of four dogs each 80%, 60%, 40%, and 20% of the trials. There are 24 permutations of four things taken four at a time, but Brogden was forced to discard 20 when deciding upon the order of presenting the various frequencies, the order being the same in each group. The permutations he chose were systematically biased, since in nine of the 12 cases a less frequent occurrence of reward followed a more frequent one. The general downward trend in frequency of reinforcement (plus the large number of trials at each level) may well have produced an adaptation or general learning effect to maintain the response at a high level with a low frequency of reinforcement.

In an early partial reinforcement experiment, Cole (**6**) conditioned 39 college students (US, shock; CS, red light; CR, eyeblink) and then extinguished the CR, after which he presented the US on alternate trials in one-half the group and gave one reinforced trial followed by three non-reinforced trials in the other half. As in Humphreys' (**29**) study, some carry-over effect from the first conditioning and extinction is to be expected.

Murphy (**49**) used a much more complicated design with adult subjects, first giving them considerable practice on a modified bagatelle (pinball) machine and then successively providing 100% "reinforcement" in the form of a flash of light, extinction, 100% reinforcement with the plunger at another setting, extinction, and two series of 20%, 10% and 5% fixed-ratio reinforcement, each pattern being followed by extinction.

As in Brogden's (**1**) study, the general sequence and the possible interactions and general learning effects as the reward progressed from 100% to 5% may possibly have acted to maintain the response at a high level even with infrequent reinforcement. The effect of the interpolated extinctions cannot be evaluated directly. Another point to be made is that the frequency of responding per minute in preliminary sessions was slightly *lower* when the reinforcement (light) was used than when it was not.

It is noteworthy that sequential effects have not been explored systematically on a large scale. What is needed is an investigation of a progressive increase in frequency of reinforcement from, say, 10% to 100% by steps of 10% and study of the consequences of the reverse sequence.

It is difficult to evaluate the influence of these sequential variables without additional experimentation. Progressive effects may well have been operating in several of the instances cited to produce considerably different results than would have been obtained with independent or matched groups.

---

[2] Apparently this fact has remained unnoticed. Hilgard (**23**) mentions "half the subjects" (p. 273) and refers to "one group" and "the second group" (p. 275).

Inter-trial interval. In addition to the usual consequences of massing or spacing practice there are two phases of the inter-trial interval that have special effects for partial reinforcement studies. Sheffield (**59**) has recently demonstrated that the spacing variable is critical in the superior resistance to extinction found after partial reinforcement. Most studies have involved massed training. Secondly, Heathers' (**20**) findings of "spontaneous alternation," recently extended by Stanley (**66**), are relevant to T-maze investigations. This factor is not crucial, but operates to retard learning if trials are massed in dual-response situations.

Number of trials and number of reinforcements. One general dilemma characterizes the partial reinforcement situation: the relationship between number of trials (or responses) and number of reinforcements. A partial reinforcement group can be matched with a continuously reinforced one on either number of trials or number of reinforcements, but not both. If, for example, 50 reinforced responses are to be given and a 50% reinforcement schedule is employed, this group will have 100 trials, and the 100% group will undergo 50 trials. The experimental question revolves around a determination of the relative effects on behavior of non-reinforced trials as contrasted with reinforced ones. One procedure is to equate on trials or responses and allow the partial group fewer reinforcements (see, for example, **12, 28, 49, 59**). If greater resistance to extinction is found for the partial subjects, it occurs despite the weighting of reinforcements in favor of the continuous group. If one is interested in acquisition, the two groups might be matched on number of reinforcements, allowing trials or responses to differ.

Brunswik (**3**) was able to circumvent this difficulty by using the correction method in the T-maze. Thus by reason of a correct response or retracing, food was received on every trial, and trials and reinforcements were equated. In a sense, however, there were no unreinforced trials. Hull and Spence (**27**) have suggested that the correction and non-correction training techniques should have different behavioral consequences. The differences in their data are not striking, and Stanley's (**66**) groups, run with the non-correction method, ranked in the same order as Brunswik's. Absolute differences may, however, exist between the results of the two procedures. In the present context it might be mentioned that Brunswik's rats received all their daily rations in the four or eight trials per day, thus providing a real possibility for gross changes in drive during a given session.

In this connection, it is possible to equate partial and continuous subjects on total *amount* of reinforcement, where trials are matched and number of reinforcements differ, by presenting twice as much food to the partial group on reinforced trials where 50% and 100% schedules are involved. Apparently this technique has not been employed.

Outside the multiple-response situation, where a correction method can be used, the very nature of partial reinforcement negates the possibility of equating both trials (or responses) and reinforcements for continuously and non-continuously reinforced groups, because in order for reinforcement to be partial some trials or responses must go unreinforced.

Another major variable to be considered is number of reinforcements per se.

Resistance to extinction has been used as a major index of the relative behavioral strengths set up by different numbers of partial and continuous reinforcements. Williams (**70**) and Perin (**51**) have demonstrated variation in extinction responses as a function of number of (continuous) reinforcements. Lloyd (**41**) has studied the same variable with periodic and aperiodic reinforcement.

These findings suggest that when a small number of reinforcements (say, less than 50 or so) are given in the bar-pressing situation, an additional variable is introduced in comparing resistance to extinction of partial and continuous groups; obviously, different tasks will have somewhat different asymptotes.

In this connection, Humphreys (**32**) presented one group of rats with 18 reinforcements on 18 bar-depressions and another group with seven reinforcements during 18 responses. Both behavior strengths were presumably increasing rapidly. Furthermore, it hardly seems likely that the consequences of non-reinforced responding in the partial group could have been very strongly conditioned with the small number of reinforcements involved (**59**).

Response strength at the initiation of extinction. The variable of strength of response at the end of learning must be considered in comparing resistance to extinction after 100% and partial reinforcement. If there is any relationship between response strength at the end of acquisition and performance in extinction, the former factor becomes critical in comparing the extinction behavior of partial and 100% groups. If on the Graham-Gagné runway, for example, the partial reinforcement group should exhibit running time several seconds faster than that of the continuously reinforced group at the end of training and then the partial group should yield greater resistance to extinction, this finding could be interpreted as a carry-over effect of behavior strength. It is not the correlation between acquisition and extinction performance that is needed, but rather the *r* between level of performance during the last few training trials and extinction behavior. Two such values have been reported (**34, 37**) for partial groups only. They consisted of a rho of .54 (S.E. =.27) and a Pearson *r* of .87(S.E. =.28) between number of non-reinforced responses in a constant period at the end of training and extinction responses. Clearly, more data on this problem are needed.

In the free-responding situation pronounced differential rates have been reported (**36, 37, 62**) at the beginning of extinction for partial as contrasted with continuous groups. At least two investigations (**32, 47**) employing the bar-pressing technique have failed to present any information on this score.

Statistical questions. In addition to the usual statistical problems of design and analysis, partial reinforcement poses one special problem that has received little attention. In several studies where enough data are available (**3, 36, 66**) a moderately high correlation between the mean and the variance seems to be present. Furthermore, there is typically a marked skewness in a given set of responses, particularly in extinction. These situations raise serious doubts as to the applicability of standard statistics for the raw data. Several methods of sidestepping these difficulties are available but have not been widely used in partial reinforcement studies. Non-normal statistics or tests involving no assumptions concerning the distribution can be employed. Another possibility that has been used (**12, 13, 66**) is a transformation of the original measures into common logarithms (for time or response scores) or angles (the arc $\sin\sqrt{\text{percentage}}$, for percentages), procedures that tend to normalize the distributions and make the variances less heterogeneous. The interested reader may consult Mueller's (**48**) recent article for a comprehensive treatment of these issues, though not in a partial reinforcement context.

A number of variables that need to be considered in order to secure clearcut results in a comparison of partial and continuous reinforcement have been indicated. Failure to control these factors is obviously not an omission peculiar to partial reinforcement studies. Experimental design in this area, however, has not been particularly rigorous. Considerably more experimentation is needed before the effects of these variables can be ignored.

III. Acquisition

In discussing the behavior of partially reinforced animals it is helpful, though perhaps somewhat arbitrary, to consider acquisition and performance separately. "Acquisition" is here viewed as improvement in executing the practiced act until an asymptote is reached, with "performance" beginning at that point. In this section we shall consider experiments wherein the former condition predominated. The first two of these, fragmentary studies conducted by Pavlov (**50**) and Wolfle (**71**), have already been mentioned and therefore need only be indicated again.

Acquisition has been studied most thoroughly in the trial situation. The present section is broken down into parts according to the methodology employed in these instances.

Classical conditioning studies. Cole (**6**), in the study previously described, was investigating the influence of sophistication (advance knowledge of the procedure) upon conditioning, so some of his subjects were informed verbally concerning certain aspects of the conditioning sequence. Cole's results show clearly that under the conditions of his experiment CRs will be maintained at a surprisingly high level for many trials even though the subject is fully informed of the order of reinforcement and non-reinforcement. His sophisticated subjects were not actually *told* to inhibit the wink on non-reinforced trials, however, so part of the continued responding can perhaps be ascribed to a misunderstanding of the experimental situation, as Cole points out.

The trend of CRs on the non-reinforced trial immediately following the reinforced one is quite alike in both groups, responses occurring less frequently then than on reinforced trials.

One of the earliest partial reinforcement experiments to arouse general interest was Humphreys' (**28**) 1939 eyelid reaction study. In acquisition, three groups of 22 college students each were trained as follows: (1) 96 trials, 100% reinforcement; (2) 96 trials, 50% semi-randomized reinforcement; and (3) like (2) except that rest intervals were substituted for the non-rewarded trials.

Humphreys equated his groups for reflex sensitivity to light and puff and reported no significant differences between them with respect to number of CRs or magnitude of response during acquisition, although Group II (50% reinforcement) was slightly lower than Groups I and III on both measures. Inspection of Humphreys' data shows that after the first few trials, frequency of CRs in the Group II curve never attains the level of the curve for Group I at any point throughout the entire 96 acquisition trials (grouped in sets of 12 trials). Thus the differences, though small, are highly consistent and suggest somewhat faster acquisition with 100% reinforcement than with 50%. The data seem out of line with Humphreys' conclusion of "presumably chance differences."

In reporting his study of the generalization of the galvanic skin response to various tonal frequencies, Humphreys (**30**) does not present data concerning progressive changes, giving only pre- and post-reinforcement results. "Acquisition" was measured by average millivolts of response following the unreinforced presentation of the four tones, twice each, in a systematically rotated order. Humphreys (**30**) uses the mean of these two values, thereby making it appear graphically that response strength was greater after 50% reinforcement than after 100%. As Fig. 1 in his later article (**31**), based on the same set of data but interpreted in connection with extinction, suggests, this is a misleading procedure, for on the *first* test (extinction) trial there was a slight superiority of the 100% reinforcement group. Thus, the data seem consistent with the position that training does not produce *greater* response strength with partial than with continuous reinforcement.

"Verbal" responses. One of the most provocative partial reinforcement studies is that of Humphreys (**29**) dealing with "expectations." The subject's task was to indicate by marking a record sheet whether or not a second light was to come on following a signal light. Seventy-eight subjects were run through the sequence: 100% reinforcement, extinction, 50% random reinforcement,[3] and extinction. With 100% reward, learning reached a 98% level. With the random series of 50% reward, learning cannot take place. There were, however, several marked and significant deviations from chance in the latter situation. These may be a function of the preliminary training in guessing flips of a coin and/or the restricted randomization of reinforcement.[4] Hum-

[3] But it was not wholly unpredictable, because "The order of 'acquisition' trials for Series II deviated from chance in that no more than two trials of a kind ever occurred together" (**29**, p. 296). This situation could have provided a discrimination for the subjects to learn. A pronounced effect of Humphreys' procedure on his extinction results has been indicated by Detambel (**8**).

[4] Humphreys' "coin flipping" series shows a trend for heads to predominate on odd-numbered trials and tails on even ones (**29**, Fig. 2). This tendency may have carried over into the 50% reinforcement training series, since the same symbols were used throughout the study.

phreys' most striking data are discussed in the extinction section of this paper.

Recently Grant and Hake (**17**) repeated Humphreys' investigation with certain modifications, including the addition of 75%, 25%, and 0% reinforcement. Their results agree with his in showing that the various groups eventually learned to respond in direct proportion to the frequency of reinforcement.

"Choice" situations. The parallel between the partial-continuous reinforcement dimension and discrimination learning is apparent. Humphreys (**30**), for example, has pointed out that early in discrimination learning reinforcement is practically on a 50% basis, while late in training when the discrimination is well formed it approaches 100%. This is by no means the complete picture, however. Obviously the discrimination learner can improve, i.e., can change the odds of the environment's paying off; in the simple case of partial reinforcement, the organism cannot alter the probability of reward feedback from the environment. In the runway situation of the Graham-Gagné variety, the partially reinforced rat must learn to respond positively to the absence of reinforcement, its concomitants and consequences. It must learn to run on non-reinforced trials as well as on reinforced ones.

Several other instances of the overlap between partial reinforcement and discrimination formation may be cited. These include Cole's (**6**) and Skinner's (**62**) fixed ratio studies, avoidance training, the use of a semi-random partial reward schedule (**29**), and dual- or multiple-response situations where one of the possibilities produces reward less than 100% of the time with the other "choices" paying off somewhere between 0% and 100%.

In the last situation, Mosteller (**44**) has pointed out the possibility of considering the consequences of the various "choices" as different in kind, each kind having a definite probability of occurring. Any kind of consequences may have a distribution of intensity (amount of food or shock) that in turn has a probability distribution varying from small to large. In the more complex instance the interactions of these parameters must be considered.

We are particularly concerned in this sub-section with situations where the experimenter provides the possibilities for responding (e.g. the T-maze) and controls the frequency of occurrence of reward for for each of the "choices." Historically, the first study of this type is Brunswik's (**3**). Large groups of rats were run on a T-Maze with different reinforcement ratios on the two arms. The ratios were: 100%:0%, 50%:0%, 75%:25%, 100%:50%, and 67%:33%.[5] (We have already mentioned certain troublesome areas of this experiment in the methodology section of this paper.)

---

[5] The three of these that contain intermediate probabilities on both sides remained unique in partial reinforcement literature until the experiments conducted by Heron (**21**), Calvin (**5**), and Stanley (**66**). Very likely such "in-between" probabilities correspond to the everyday situation more than ratios like 50:0, in which only one alternative is "am-

All patterns except 67:33 resulted in definite learning by the end of the 24 trials. The means and variances are highly correlated, with each of the distributions of errors markedly skewed in a positive direction. In the present context it is noteworthy that the 100:0 rats made fewer errors than the 50:0 animals, although the difference does not quite reach the 5% level of significance.

Instead of extinction Brunswik (**3**) reversed the training of his various groups. Again the 100:0 rats learned most readily; only the 75:25 and 67:33 groups seemed unable to relearn. Humphreys (**32**) interprets these data as extinction, but the operations differ from those usually employed in extinction. Detambel's (**8**) findings, discussed in a later section, suggest that in a different situation the two methods do not yield comparable results.

Brunswik's (**3**) conclusion that response strength in his situation is a compromise between the influence of probability difference and probability ratio does not seem to be a complete statement. The assumption that the 100:0 and 50:0 ratios are identical except for the probability differences is questionable when the "all-or-noneness" of the former is contrasted with the half-time reward on one side of the latter. Similarly, although the ratios 50:0, 75:25, and 100:50 have equal probability differences, there is a confounding of partial reinforcement on the favorable side, on both sides, and on the less favorable side that might be used to explain the progressive increase in errors. The same kind of qualms apply to 100:50 and $67\frac{2}{3}:33\frac{1}{3}$, where the odds in each instance are 2:1. In any event, there is no reason to suppose that any elementary mathematical function of reinforcement percentages will describe the effect.

A detailed analysis of this sort of data has not been performed in terms of the algebraic summation of tendencies strengthened by reinforcement and weakened by its absence, similar to Spence's (**65**) formulation. The additional feature that would have to be included is the absence of 100% reinforcement even for a "correct" response. One mathematical formulation is underway (**44**) on this problem.

Stanley (**66**) ran five squads of seven matched rats on a T-unit with these ratios: 100:0, 50:0, 75:25, 100:100, and 50:50. A criterion of learning was employed such that 13 days of eight trials each were involved on the average. These were followed by extinction trials. The non-correction method was used. His design permitted analysis of the total variance or covariance into three parts, that attributable to blocks, reinforcement ratios, and error. Number of reinforcements

---

biguous." The analogy of these ratios to certain of Lewin's (**40**) types of conflict is apparent, e.g., 100%:50% and approach-approach.

were matched within each *block*, being based on the number received by the 100:0 rat in that block. Thus, number of trials varied systematically from group to group.

In terms of errors, the ranking of the three groups was similar to Brunswik's (**3**): 75:25 most, 50:0 next, and 100:0 least. For these three groups speed of learning was positively related to number of reinforcements not merely trials, as Humphreys (**32**) has contended. With reinforcements equated, the 100:0 and 50:0 groups did not differ significantly in regard to wrong turns *per trial*, but the 75:25 animals were distinctly more erratic than either.

Denny (**7**) ran 100:0 and 50:0 groups on a T-maze to test his hypothesis that the almost equal learning level reported by Humphreys (**32**) for partially and continuously reinforced groups when number of trials was equated but reinforcements favored the 100% group could be attributed to secondary reinforcement from the goal box on the non-reinforced trials.

Denny interprets his acquisition findings as supporting the secondary reinforcement hypothesis. Using the initial trial of each day, he shows that when secondary reinforcement is allowed to operate (control condition) acquisition proceeds almost as well under 50% as under 100% reinforcement. Also, the expected (and significant) difference in favor of continuous reinforcement is found when secondary reinforcement is minimized. If the first *two* daily trials are used for the control groups, however, the superiority of the 100% rats reaches the 10% level of significance. This finding suggests that the 100% control group may have learned more rapidly than the 50% control group.

The interaction of partial and secondary reinforcement has been the topic of several other investigations of the choice variety. Ehrenfreund (**10**) has shown that the secondary reinforcing properties of goal boxes are significantly weaker when previously associated with partial (50%) reward than after the same number of continuously rewarded trials. The number of primary reinforcements was 40 for the 100% group and 20 for the 50% rats.

In a similar study, Saltzman (**55**) found superior learning for his partially reinforced rats when working for the secondary reward cues. The 100% group received 25 rewarded runs; the partial animals had the same training plus 14 interpolated non-reinforced runs. In a second study interspersed trials on the original runway produced equal or superior performance for the 100% group.

The number of reinforcements employed in both these studies is noteworthy, since a similar investigation (**19**) has shown an increased tendency to learn to run to a secondarily rewarding goal box as the number of continuous, primary reinforcements in training increased from 25 to 50 to 75. Whether or not this parameter acts differentially in the partial-continuous context is an experimental question.

In this connection a recent study (**14**) showed that rats will press a bar more frequently for a secondary reinforcing stimulus when the acquired reward property stems from previous pairing of the stimulus with aperiodic as contrasted with continuous primary reinforcement. The mean number of bar-pressing responses in 45 minutes was 39 for the 100% group and 57 for the aperiodic group. The difference just fails to reach the 5% level of significance; the number of cases was small, four in the 100% group and five in the aperiodic squad.

Employing various combinations of 10 electric light bulbs on a vertical board, Calvin (**5**) studied the written predictions of college students as to the occurrence or non-occurrence of a buzzer following each combination. One of the combinations was succeeded by the buzzer a greater or lesser proportion of trials.

Calvin concluded that learning in this situation depends on the probability of reinforcement for the particular pattern and also on the probabilities characteristic of the other stimulus compounds.

Only Heron (**21**) has conducted an investigation in which rats reinforced periodically with respect to time were presented with alternatives set up by the experimenter. After preliminary training of 48 rats to press a single lever in order to receive a pellet of food and then to operate two levers, he set the right lever to "pay off" a pellet every three minutes and the left one each six minutes. Under these conditions the frequency of responding was about the same on the two levers until the seventh hour, when it increased on the right lever. Thereafter, frequency was maintained at a relatively high level on the right lever and decreased slightly and gradually on the left one to the end of the 17 one-hour periods. The ratio of responses on the two levers was about what is to be expected from Skinner's (**62**) essentially linear (negative) relationship between frequency of responses and size of time interval between reinforcements under the periodic regimen.

Straight runway. The two experiments that Finger (**12, 13**) reported in 1942 have been challenged by Lawrence and Miller (**39**) on the grounds that interchanging the goal and starting boxes, with the rat still inside, eliminated distinctive cues for the incompatible responses of searching for food and running.

With regard to acquisition under a partial reinforcement schedule, Finger's results are fairly clear. His partial reinforcement groups were all reliably slower than the 100% animals. Finger's data on acquisition are generally in line with other results in showing somewhat more rapid learning with continuous reinforcement than with partial. It does not seem that the lack of distinctive cues was acting differentially, although its possible confounding effects cannot be partialed out. By analogy to the Lawrence-Miller argument, the learning of the partially reinforced rats may have been somewhat retarded on the trials following non-reinforced runs since there would be conflict between the tendencies

TABLE II

A Comparison of Acquisition Under Continuous and Partial Reinfopcement

The heading "Direction of Difference" means that the group listed under "1" learned more readily than the group given under "2." A blank space under "*P*-Level" indicates that the probability of the difference did not reach the 5% level of confidence. An asterisk at any point in the body of the table refers to data estimated or calculated by the present writers.

Investigator and Situation	Direction of Difference		Ratio of Smaller to Larger	*P* Level	Comment
	1	2			
Brunswik (**3**) T-maze	100:0	50:0	.45		
	100:0	75:25	.35	5%	
	100:0	100:50	.25	1%	
	100:0	67:33	.15	1%	
Denny (**7**) T-maze	100:0	50:0	.96		Control
	100:0	50:0	.69	1%	Experimental
Finger (**12**) Runway	100	50	.65*	1%	
Finger (**13**) Runway	100	50	.82	1%*	Common logs
Grant & Hake (**17**) Verbal Responses	100	25, 50, 75	—	1%	Abstract
Humphreys (**28**) Eyelid	100	50	.97		Frequency and magnitude
F. Sheffield & Temmer (**58**) Escape-avoidance	Escape	Avoid.	.46	1%	
V. Sheffield (**59**) Runway	100	50	.99*		
Stanley (**66**) T-maze	100:0	50:0	.59	1%	Errors
	100:0	75:25	.43	1%	Errors
	100:0	50:0	.54	1%	Trials to criterion
	100:0	75:25	.57	1%	

to look for food and to run. There is a slight indication of this influence in Finger's data. The effect would be diluted, of course, by learning.

Though presumably she controlled "secondary reinforcement" somewhat by removing the food cup on non-rewarded trials in her runway situation, Sheffield (**59**) found only a very slight acquisition difference favoring the 100% group over the 50% subjects ($P=.30$). This finding held for both massed and distributed practice, with inter-trial intervals approximately 15 seconds and 15 minutes, respectively.

Escape and avoidance training. In a recent investigation that has broad theoretical implications, Sheffield and Temmer (**58**) found that rats learned more readily to escape from a shock situation than to avoid it. The learning curves rose together, after which the curve for avoidance training dropped gradually to a lower level. Since avoidance means a lessening of reinforcement, while escape involves 100% reinforcement, these results are in line with other acquisition data discussed in this section.

Tabular Comparisons. Table II contains comparative data from the relevant experiments referred to above. In not a single instance did subjects learn better under partial reinforcement than under continuous. Only Denny (**7**), Humphreys (**28**), and Sheffield (**59**) failed to find significant superiority for 100% reinforcement in *any* phase of their studies.

Over-all, it can be said with considerable assurance that *acquisition proceeds somewhat more rapidly and reaches a higher final training level under continuous reinforcement than under partial reinforcement.* With prolonged training partially reinforced groups may approach the same level of proficiency attained earlier in training by continuously reinforced subjects.

IV. PERFORMANCE, MAINTENANCE AND RETENTION

In this paper a distinction that is somewhat arbitrary has been drawn between acquisition and performance. By acquisition we have meant the progressive changes in behavior prior to the stabilization of responding at a more or less constant level. In the bar-pressing apparatus, for example, it is frequently difficult to determine when acquisition occurs. One response of the specified variety (magnitude, direction, etc.) may be a sufficient (as well as a necessary) condition for starting the behavior off at a relatively constant and stable level. Certainly for practical purposes one can distinguish learning from already established behavior. In this section we are concerned with the role played by partial reinforcement in maintaining behavior previously acquired.

One of the classical studies in this area has been reported by Skinner (**62**, p. 127 ff.). After previous conditioning and participation in other experimental activities, four rats were run in a rotated order under several different periodic reinforcement regimes: three, five, seven, and nine minutes. Each rat underwent two one-hour periods on successive days at each interval. The mean number of responses per hour were, for the four periodic intervals respectively: 319, 266, 220 and 169. The relationship is essentially linear over the range of four points involved.

A second kind of partial schedule, called fixed ratio reinforcement, was also developed and studied by Skinner (62). The operations defining it consist of presenting a reward for a fixed number of responses. Thus a rat might be reinforced after every 20th bar-depression. Skinner's basic experiment (**62**, p. 286 ff.) consisted of running six previously conditioned rats through a random order of fixed ratios: 48, 96, or 192 responses for one reinforcement. The characteristics of the curves are striking: (1) Just after receiving food for a response there is a pause that increases as the ratio decreases. (2) There follows a period of gradually increasing responding that (3) seems to reach a stable state shortly before the next reinforcement. (4) The over-all slopes of the cumulative response curves appear to steepen as the number of responses per reinforcement increases, but the effect is not pronounced. This finding may be related to the fact that a greater portion of the time is devoted to eating at the larger ratios. In an hour almost three times as many pellets are received in the 48:1 case as contrasted with the 192:1 regime.

Eating time may be somewhat compensated for by the longer pause after eating at the lower ratios, but the relationship between ratio and rate is by no means a simple one such as that holding between rate of responding and length of time interval under a periodic reinforcement schedule.

It should be noted that the behavioral effects of periodic and fixed ratio reinforcement are opposed. In the first procedure, increasing the time interval between reinforcements produces a decreased rate of responding. Under fixed ratio reinforcement, increasing the number of responses per reinforcement tends to yield a higher rate overall. In addition, the qualitative characteristics of the curves are markedly different. Skinner's cumulative response curves under periodic reinforcement are smooth, lacking the discrete, step-wise nature of the fixed ratio curves. Fixed ratio reinforcement strengthens rapid responding; periodic reward ultimately strengthens responding.

Chronologically, the next systematic investigation of performance involving partial reinforcement was Brodgen's (1). Three groups of four dogs were employed. One group had leg flexion conditioned to bell with the US of shock. The second group had similar initial training to a high level of performance followed by the substitution of food for shock. The third group underwent a classical conditioning routine of salivation to the CS of bell and the US of food. After conditioning had attained a 100% frequency for all three groups, with continuous reinforcement, partial reinforcement was instituted. This consisted of 100 trials under each of the following frequencies: 80%, 60%, 40%, and 20%, with the preceding schedule of 100% available as the continuous-reinforcement baseline.

In all three groups frequency of CRs with 20% reinforcement was less than that of the 100% condition—significantly so in all but the shock-flexion group. The latter animals showed a progressive *increase* in frequency of CRs as reinforcement decreased from 100% to 40% followed by a drop in CRs at 20%. The other two groups yielded a more or less progressive decline in frequency as reinforcement decreased. The lowest mean frequency of response attained was above 80% for all three groups. Latency and magnitude of conditioned salivation yielded generally decreasing trends as frequency of reinforcement decreased from 100% to 20%.

We have commented in Section II on the possibility that the progressive error inherent in Brogden's design and the degree of overlearning employed operated to maintain the response at high frequency throughout and thereby reduced any decremental effects of the lower frequencies of reinforcement.

The poorer conditioning with unavoidable shock (100% reinforcement) in Brogden's study is consistent with other findings (**24**, **57**). As Sheffield (**57**) has suggested, one possible action of unavoidable shock is to condition to the experimental situation other behaviors (or "states" producing them) that are incompatible with the response studied. Brogden himself reports decreased "vocal response, diffuse behavior, and general restlessness" (**1**, p. 422 f.) with decreased frequency of (shock) reinforcement.

The conclusion from Brogden's data is clear: Low frequency of partial reinforcement produces a response decrement.

Murphy (**49**) concluded that there was little relationship between rate of responding and frequency of reinforcement per response in his pinball situation. The mean number of responses per minute for continuous reinforcement, 1/5, 1/10, and 1/20 were: 17.4, 15.2, 15.6, and 16.4. (These data have been averaged by the present writers from Murphy's several conditions.) In 70% of the measurements, 100% reinforcement yielded rates above those for the fixed ratios. It seems that human subjects will continue to work almost as hard for infrequent "reinforcements" as for frequent ones of the kind Murphy used. The specific action of the variables of sequence and number of reinforcements, previously noted, cannot be partialed out. It is surprising that the subjects did not verbalize the situation and speed up their responses if they were really "working for" the light. It is noteworthy in this connection that the responses per minute when the light-reinforcement was given either continuously or intermittently is slightly *lower* than in preliminary practice periods at the beginning of each session when the subjects recieved no light and were instructed to operate the plunger at a "natural or comfortable" pace and force. Apparently "reinforcement" of the light added little or nothing to the instructions as far as number of responses per minute was concerned.

Where the organism is free to respond, as in situations involving bar-pressing or a similar instrumental act, it is difficult to isolate an acquisition phase. Thus these studies bear mostly on performance during conditioning. Two investigations, one involving periodic reinforcement (**37**) and the other aperiodic (**36**), may be mentioned. In the first instance the drive involved was thirst and the response was bar-pressing by rats. Performance, measured in terms of rate of responding, was significantly higher under continuous reinforcement than under either a one-minute or a two-minute periodic schedule. This finding bears particularly on the question of response strength at the beginning of extinction. One might question whether differences in resistance to extinction—favoring groups partially reinforced in conditioning—are attributable merely to a carry-over of greater response strength. In these data, the situation was clearly and significantly reversed, so extinction differences favoring intermittent reinforcement cannot be accounted for by differential rates of response at the end of conditioning.

The second study (**36**), involving aperiodic reinforcement, raises a related problem. The pecking response of pigeons was employed with hunger as the drive. The aperiodic schedule averaged one reinforcement every minute. Rate of responding was approximately five times as high for the aperiodically reinforced birds as for the continuously reinforced ones in a constant period of time just prior to extinction. This comparison is not a fair one, however, since about one-half of the time was devoted to eating in the case of the continuous schedule and only about 10% for the aperiodic one. Even when the groups are matched on this score, though, the rate of responding was more than twice as high for the partial reinforcement subjects. Differences in extinction were in the same direction and of even greater magnitude.

In any event, the bar-pressing study cited previously (**37**) makes it clear that response strength under performance conditions is not the only variable operating to produce differential extinction effects after partial and continuous reinforcement.

In the present performance context might be mentioned Humphreys' (**30**) finding that generalization of GSR to tones as far removed as 26 *j.n.d.*'s from the reinforced tone takes place much more readily following partial reinforcement than after continuous reinforcement.

Mosteller and Nogee (**45**) have recently studied gambling behavior in a situation similar to partial secondary reinforcement. The subject is shown a poker hand and is told what the (statistically) "fair" odds are. He either bets or does not bet a nickel against certain odds of drawing a hand exceeding the one shown to him. Latency of response is one measurement made. If a bet is taken the hand is dealt and money is paid or received. By varying the betting odds on a particular hand the point is found at which a given subject takes the bet 50% of the time. By using a number of different hands, a series of such indifference points are found. These are used along with a theoretical equation to obtain a utility curve for money. The data were still under analysis at the time of writing, but the trends suggested that highly motivated (poorer economically) subjects were rather consistent about taking unfair bets, while less motivated subjects consistently wanted better than fair odds.

Skinner (**63**) has recently found that the pecking response of pigeons is maintained at a relatively constant rate during the major portion of an eight-hour performance session. The aperiodic schedule averaged a reinforcement once every five minutes, and during the period around 50,000 responses were given. There appeared to be only a moderate decline in frequency of responding toward the end of the session.

One systematic study of long-term retention might be mentioned. Skinner (**63**) has tested small squads of pigeons on a partial reinforcement schedule after intervals of two and four years of no experimental activity. There appears to be little if any loss or change in the response.

Partial reinforcement procedures—particularly periodic and aperiodic—have served as valuable baselines in the study of the effects of several variables on behavior. The stability of responding under these regimes provides a sensitive measure of the action of such factors. An additional advantage of these partial reinforcement schedules is the opportunity for relatively prolonged single experimental sessions by way of staving off satiation effects. Some of the variables that have been studied under these conditions include drugs such as benzedrine, phenobarbital, insulin, and caffeine (**62, 64, 69**); amount of reinforcement (**35, 62**); drive (**22, 34, 62**); and quantitative variations in the response (**62**).

Empirical principle. The data seem to be in condition to derive an empirical principle that is, of course, open to further experimental test. A number of variables are operative both in experimental and every-day situations that will cut across the present generalization. The amount of reinforcement may, for example, be critical. It seems questionable that without an exceptionally high drive human beings would work as hard for exceedingly small amounts of money administered after every response as they would for large amounts given occasionally. In any event, the data concerning maintenance of behavior under partial reinforcement are fairly clearcut. The following characteristic of performance under a partial reinforcement schedule emerges: *All other things equal, performance under a partial reinforcement schedule tends to be somewhat lower than that under a continuous reinforcement one* *as measured in terms of single responses.* The behavior is stable under a partial reinforcement schedule. Instances can be found indicating either higher or lower response strength in the partial instance, but, at least for small frequencies of reinforcement, the typical response strength is lower with partial reinforcement than with continuous reinforcement.

One exception to this principle may be indicated, namely, fixed ratio reinforcement in the free-responding situation. Under this schedule there is some increase in rate of responding as frequency of reinforcement is decreased, i.e., as number of responses per reinforcement is increased. As Skinner (**62**, pp. 274-300) has pointed out, this procedure involves a special discrimination in which rapid responding is strengthened since it brings reward sooner. The effect is not as pronounced as the decrease in rate found with increased time interval in periodic reinforcement. Both the fixed ratio and random ratio reinforcement schedules need more investigation. This question is discussed further in Section VI.

The question of practical vs. statistical significance arises in connection with the lower response strength under a partial reinforcement regime. In everyday behavior it does not seem of great consequence whether partial reinforcement produces a stable response level below that of continuous reinforcement at a statistically acceptable level (as is indicated by several studies). The important point is usually for occasional responses to occur. Partial reinforcement unquestionably accomplishes this.

V. EXTINCTION

Extinction has been a ubiquitous measure in the comparison of partial and continuous reinforcement. A large number of studies have employed this criterion of response strength in comparing the two conditions of reward and in evaluating the influence of partial reinforcement on behavior. The maintenance of behavior in the absence of reward after a partial reinforcement training regime raises several theoretical problems that are treated in a later section. The practical implications are far-reaching. It can be said that extinction has been the main mirror reflecting response strength in the partial reinforcement situation.

The data regarding the comparative effects of partial and continuous reinforcement on extinction are summarized in Table III. It includes all the studies the authors could search out that involved the partial-continuous comparison. Like any such summary it is incomplete and needs elaboration, but the trend of the findings from a number of different experimental situations is clearly represented.

TABLE III

A Comparison of Resistance to Extinction After Continuous and Partial Reinforcement

The heading "Direction of Difference" means that the group listed under "1" yielded greater resistance to extinction than the group given under "2." A blank space under the heading "*P*-Level" means that the difference did not reach the 5% level of confidence. A dash indicates that the data were not given in the original report. An asterisk refers to data calculated or estimated by the present writers. Consistent data for more than one comparison have been combined.

Investigator and Situation	Direction of Difference		Ratio of Smaller to Larger	*P* Level	Comment
	1	2			
Denny **(7)** T-maze	100:0	50:0	.96		
Finger **(12)** Runway	50	100	.56	2%	
Finger **(13)** Runway	100	50	.97		
Gilinsky & Stewart **(15)** Learning Mazes	50	100	—	—	Abstract
Grant & Hake **(17)** Verbal Responses	25, 50, 75	100	—	1%	Abstract
Grosslight & Child **(18)** Plunger-pull	Partial	100	.49, .10	(6%), 1%	No. of S's, trials
Humphreys **(28)** Eyelid	50	100	.47	1%–5%*	Frequency & magnitude
Humphreys **(29)** Verbal Responses	50	100	.55*	1%*	
Humphreys **(31)** GSR	50	100	.68*	1%	
Humphreys **(32)** Bar-pressing	50	100	.71	1%*	Equal no. of reinforcements
Jenkins *et al.* **(36)** Key-pecking	Aperiodic	100	.26	1%	
Jenkins & Rigby **(37)** Bar-pressing	Periodic	100	.48	2%–4%	Equal no. of reinforcements
Mowrer & Jones **(47)** Bar-pressing	25, 33, 40, 50	100	.60	1%–2%	
Murphy **(49)** Pinball	5, 10, 20	100	.31*	1%*	
F. Sheffield & Temmer **(58)** Escape-avoidance	Avoid.	Escape	.47	1%	
V. Sheffield **(59)** Runway	50 100	100 50	.76 .95	1%–3%	Massed Spaced
Stanley **(66)** T-maze	Partial Partial	100 100	.36 .57	1% 1%–3%	Trials Errors

The consistency of the data in this table is striking. Of the 17 separate studies listed, only three failed entirely or in part to find extinction differences significant at the 5% level of confidence favoring the groups partially reinforced in training over the continuously rewarded ones. About 35 comparisons are condensed in Table III. Only four of the separate measures yielded differences in a direction opposed to this trend, and in none of the three experiments concerned did the differences reach the 10% level of confidence. All three have special aspects and are discussed in the following paragraphs.

The first exception to the noted trend is in Denny's (**7**) study, where slight differences favoring the continuous groups were found in resistance to extinction for experimental subjects for which secondary reinforcement had been minimized. In the control groups the results were similar, but again not significant.

The action of an important variable in Denny's study has been demonstrated by Sheffield (**59**). She found that resistance to extinction was slightly (but not significantly) better after continuous reinforcement when the training trials had been *spaced*. With *massed* practice the findings were reversed and the typical difference favoring the partial group was statistically acceptable. The spacing variable certainly affected Denny's results, since he interpolated intervals from 20 to 30 minutes between training trials where Sheffields' spaced condition was only 15 minutes. The theory behind the Sheffield experiment is discussed in detail in Section VII.

The fourth exception to the trend in Table III (again not attaining statistical significance) is found in the second of two studies by Finger (**13**). The continuously reinforced groups exceeded the partial subjects in trials to reach a criterion of non-response by a very small (and insignificant) margin. The action of lack of distinctive cues inherent in Finger's procedure of employing the goal and starting boxes interchangeably has already been discussed. The superiority of the continuous animals in Finger's second study is of special note since the findings were the reverse of those in his first investigation. The major difference between the two experiments seems to be the interpolation of a 24-hour interval between acquisition and extinction in the second case. Since the difference in the second study was minute and unreliable, this finding could be attributed to sampling. In the first study the difference in the opposite direction was significant at better than the 2% level of confidence. It is noteworthy that the partial animals were slightly faster starters on the average, but slower runners in the extinction of the second study.

In any event, Finger's data cannot be considered a serious contradiction of the overall principle, since the interaction effects of partial and derived reinforcement are not determinable without further investigation.

One other study not included in Table III requires comment, namely, the investigation by Keller (**38**) of the effects of the sequence of partial and continuous reinforcement. The methodological import of these data has already been pointed out. The post-continuous group significantly exceeded the post-periodic group in responses in the first five minutes of extinction; the responses of the partial animals exceeded those of the continuous ones in the second and third hours. After three hours the post-periodic group showed a very slight superiority (1%). These findings are consistent with those from other studies (**37**, **62**) showing more rapid responding early in post-continuous extinction than in post-periodic extinction. The fact stands out, however, that when 30 continuously reinforced responses were followed by 30 periodic ones or vice versa, the asymptote in three hours of extinction was inappreciably different in the two cases. Recent additional evidence (**9**) on the continuous-partial sequence supports Keller's findings, except that the rats starting extinction after continuous reinforcement slightly exceeded the other group in total extinction responses (ratio of 1.2). A special aspect of this investigation was that initial conditioning with 20 continuous reinforcements and one hour of extinction preceded the initiation of the experiment proper.

In both of these sequential studies the question of the effects of larger numbers of reinforcements is still open to experimentation.

There remain several other studies and points that need treatment. The earliest extinction comparison of periodic and continuous reinforcement has been reported by Skinner (**62**). Although not providing numerical information, he notes that the post-periodic cumulative extinction curve in the bar-pressing situation as compared with the post-continuous one rises more slowly, reaches a higher asymptote (with prolonged extinction), falls off more slowly, and lacks the cyclic fluctuations which are attributed to emotional effects in the continuous case. He concludes that "no amount of continuous reconditioning will yield an extinction curve of the height obtained through even small amounts of periodic reconditioning" (p. 138).

Further support for these statements may be found in a recent investigation (**37**). It was shown that 90 reinforcements administered according to a two-minute periodic schedule yielded an appreciably greater number of extinction responses than did 2400 continuously reinforced responses (ratio of 1.4). The continuously reinforced subjects had the advantage in both number of reinforcements and total number of conditioned responses as shown by ratios of about 27 to 1 and 5 to 1, respectively.

As regards extinction after fixed-ratio reinforcement, Skinner's curves show exceedingly high rates of responding, from which he concludes that "if a response is going to be emitted at all, it will come out as soon as possible" (**62**, p. 295). The rate curves in extinction are characterized by steep slopes and abrupt cessation of responding. Data comparing extinction after a wide range of fixed ratio reinforcements are not available, but the findings of Mowrer and Jones (**47**), who used ratios from 1/1 up to 1/4, are contained in Table III. Even with this narrow range they found large differences in single extinction responses favoring the smaller fractions.[6]

Several studies condensed in Table III do not provide information on differential responses strengths at the initiation of extinction. Mowrer and Jones (**47**), for example, found that resistance to extinction varied inversely with the size of the fixed ratio reinforcement, but they did not report response strength at the end of conditioning. From other data (**62**) it would appear that higher fixed ratios are characterized by higher rates of responding.

One special point may be made here. Resistance to extinction is usually *greater* after partial reinforcement, and typically, response strength in conditioning is *less* for the same case. For example, periodic reinforcement yields a lower rate of responding in conditioning than does continuous reinforcement (at least when the periodic interval is fairly long); this lower rate appears to carry over into extinction. When the latter is prolonged, the periodically reinforced subjects ultimately exceed the continuous ones in total number of responses, although rate of response early in extinction may be greater for the continuous subjects (**37**, **38**, **62**). With a truncated extinction, however, the situation may be reversed. When the higher response strength characterizes the partial reinforcement instance in conditioning (as in fixed ratio situations), the potential effect of this variable cannot be ignored. From a practical standpoint, as has been pointed out, the relative response strengths in performance may be of little consequence as long as the desired behavior occurs occasionally.

Further questions are raised by the rather small number of reinforcements administered by several investigators (**7**, **12**, **13**, **32**), a point previously treated. It will be noted that the comparison in Table III for Humphreys' (**32**) bar-pressing study is based on groups with equal numbers of reinforced responses. It will be recalled that two groups received 18 reinforcements, in the one case in 52 responses and in the other in 18. The other two groups were reinforced seven times in either eighteen or seven responses. In a comparison of the two groups matched on total number of trials (18 responses), the continuously reinforced rats were slightly superior in extinction behavior (by an average margin of less than one response), with the group averages around 25. But they also had two and one-half times as many reinforcements. Increasing resistance to extinction with increasing numbers of reinforced responses in conditioning is a point previously discussed in detail. In this case there is clearly a confounding of the contribution of number and pattern of reinforcements. For this reason, these data were omitted from the table. The unknown but possibly differential effects of small numbers of reinforced responses across the partial-continuous variable have been suggested previously.

---

[6] The abscissa in Figure 3 of the Mowrer-Jones (**47**) article should be percentage of total number of bar-pressings reinforced during final training and the ordinate, correct as shown, the total number of bar-pressings during extinction. The abscissa actually consists of the five reinforcement ratios used in the experiment: 1/1 (ratio of one reward for one lever-pressing, or 100% reinforcement); 1/2 or 50%; 1/2.5 or 40% (randomly administered); 1/3 or 33 1/3%; and 1/4 or 25%. The authors have plotted the *reciprocals* of these ratios. Replotting the graph with the percentages given above fails to produce even an approximation to the linearity claimed by the investigators.

Detambel (**8**) has put forward a cogent criticism of Humphreys' (**29**) study in which subjects were to guess whether a second light was to come on following a signal light. The problem is posed by the fact that Humphreys' acquisition curves begin near 50%, as would be expected from the instructions, including preliminary exposure to guessing flips of a coin, but in extinction the curves for both conditions of reinforcement (50% and 100%) drop in a few trials to a near-zero level. According to Detambel's analysis, Humphreys' instructions set the stage so that whenever the subject's "yes" was followed by the second light the "yes"-response was strengthened. Whenever the second light failed to occur following a "no"-response, this alternative behavior was reinforced. Thus in extinction the complete absence of the second light strengthened the "no"-response to such an extent that it reached a level near 100% frequency, while the competing response of "yes" was never reinforced and approached zero.

Detambel tested his hypothesis in a situation consisting of a pair of keys and one light. A correct response was a key closure that turned on the light. The subjects were instructed to press the key they thought would turn on the light. Two groups parallel to Humphreys' were run (100% and 50% reinforcement, with Key A correct in training and Key B correct in extinction) along with a 50% and a 100% reinforcement group exposed to no reinforcement in extinction. A minimum of 120 extinction trials was given, with 25 subjects in each of the four groups.

When the procedure resembled Humphreys', the extinction curves were similar to his. Striking differences were found when reinforcement of the incompatible response was omitted in extinction; both curves dropped towards a chance level (50%), with a faster approach following 100% reinforcement. These findings clearly support Detambel's contention that Humphreys' procedure was contaminated by the introduction of an incompatible response so that extinction consisted of the differential weakening of one response and the strengthening of the other.[7]

---

[7] We have not intended to imply support for either an interference or an adaptation theory of extinction in this and related contexts. Extinction may well involve the learn-

We have already indicated the difficulty of evaluating the effect on Humphreys' results of running the same group of subjects through the sequence: conditioning with 100% reinforcement, extinction, reconditioning with 50% reinforcement, and re-extinction.

Using independent groups, Grant and Hake (**17**) have enlarged on Humphreys' findings by employing 100%, 75%, 50%, 25%, and 0% reinforcement. Their data show greater resistance to extinction after the partial schedules.

There still remain several studies dealing with the consequences of partial reinforcement for resistance to extinction. These have not been included in Table III either because they did not involve a continuous-reinforcement baseline or because they employed special reinforcing stimuli.

The first of these is a systematic investigation by Estes (**11**) of the effects of shock on behavior in the bar-pressing situation. Estes' administration of periodic shock was a special case of the procedure as defined by Skinner (**62**), since the former's involved presentation of shock for all responses within a portion of the total exposure time. Thus in one case (Experiment K) for one minute of every four all bar-depressions were followed by shock, while during the remaining three minutes responses were neither rewarded nor punished. Straight extinction followed this treatment. Control groups received shock for every response or extinction in a similar period. Two separate studies yielded comparable data. The initial effect of periodic shock was less than that of shock for every response, i.e., more responses were given under the periodic shock condition. In extinction following these treatments, however, the number of responses was consistently smaller for the periodically shocked groups than for those given shock for every response. About 50% more responses were given in extinction after continuous shock in one experiment and 15% in the other.

In his two-bar situation previously described, Heron (**21**) gave 12 days of extinction to one-half of the rats on the right lever and to the other half on the left lever with the other one yielding reinforcement once every three minutes. Responses in extinction were more frequent on the formerly three-minute reinforcement lever for the first four days.

---

ing of incompatible responses or of not-responding. The subject is certainly doing something else when it fails to respond in extinction. Rather, we have intended to stay close to the common usage implied by the term extinction: omission of the reinforcing stimulus and the decreasing response strength that accompanies this operation. Our interpretation is that Humphreys (**29**) did not follow this model, since his extinction procedures involved the strengthening of an incompatible response that (if it existed at any strength at the initiation of training) was presumably weakened during the conditioning of the alternative behavior. Brunswik's (**3**) relearning situation was omitted on similar grounds. Detambel's (**8**) experimental analysis indicated that in Humphreys' case the two procedures lead to very different consequences.

On the fifth and sixth days the average number of responses on the two levers was about the same. On days 7–12, more responses were given on the lever formerly producing reinforcement once every six minutes. These findings are consistent with other data (**37, 38, 62**) in showing that the low rate for the low frequency of reinforcement in conditioning is maintained in post-periodic extinction so that if the latter is prolonged the ultimate number of responses exceeds that for the group (or lever) more frequently rewarded in training.

It should be noted that Heron's study does not involve a simple case of extinction; responding on the non-reinforcing lever was presumably maintained to some extent by three-minute periodic reinforcement on the other lever. Extinction effects generalized in the opposite direction: The number of responses on the lever paying off every three minutes dropped. The former effect may account for the relatively high rate of responding on the "extinction" lever even after 12 extinction periods. Also, the fact that for half the subjects extinction involved the additional shift from six- to three-minute periodic reinforcement on the left lever must be considered.

Another study relating partial reinforcement to extinction was performed by Cole (**6**). Fixed-ratio reinforcement was employed in conditioning either every other trial or every fourth one. The frequency of CRs in extinction after 50% reinforcement followed a roughly stepwise function although some CRs appeared on all early trials. In this connection, it should be noted that the discrimination of blinking on alternate trials was not particularly well established when extinction was instituted even though some of the subjects were explicitly instructed regarding the nature of the treatment.

Eight subjects were reinforced on the 1st, 5th, 9th, etc. trial. In extinction the effect of this training was very pronounced. Except for one set of points (in 28 trials) fairly late in extinction, pronounced peaks in frequency of CRs appeared exactly every fourth trial starting with the first extinction trial. Frequency of CRs decreased to a minimum by the third or fourth extinction trial of a given block and rose greatly on the first trial of the next block. The orderliness of the data is striking. Overall, of course, the general trend was towards decreasing frequency of CRs. It should be noted that the eight subjects were fully acquainted with the procedure except that they were not instructed as to the beginning of the extinction series. It is difficult to evaluate the effects of the earlier 100% reinforcement and extinction on the later training.

A comparison of periodic and aperiodic reinforcement in resistance to extinction has recently been performed by Lloyd (**41**). The total number of responses for groups given larger numbers of reinforcements was about the same in four one-hour periods of extinction, although the post-periodic rats produced significantly more bar-pressing responses (at the 5% level) in the first extinction hour.

Finally, an empirical generalization following from the material of this section can be stated. It is quite straightforward in view of the consistency of the data: *All other things equal, resistance to extinction after partial reinforcement is greater than that after continuous reinforcement when behavior strength is measured in terms of single responses.* A necessary condition for the operation of this principle appears to be massing of training trials or responses, at least with infra-human subjects.

The reasons for qualifications concerning measures in terms of single responses and massing of training trials are set forth in the following section on theory. The explanation of the greater resistance to extinction of behavior based on a partial as compared with a continuous reinforcement schedule is also deferred to the theory section.

The practical question of how best to build a response to withstand elimination in the absence of primary reinforcement is clearly answered by these data: *Administer the reinforcements according to a partial regime.*

VI. THEORY

Explanations of partial reinforcement phenomena fall into two mutually exclusive categories: (1) those based on aspects of stimulus-response learning theory and deductions therefrom; and (2) positions referring to cognitive processes and the acquisition of expectations and hypotheses. This breakdown will be followed in the present section.

The facts deriving from a comparison of partial and continuous reinforcement that a complete theoretical treatment must account for are these: (1) more rapid acquisition with 100% reinforcement, (2) a somewhat higher level of performance with 100% reinforcement, and (3) considerably greater resistance to extinction after partial reinforcement.

S-R learning theory. A straight S-R interpretation, that reward strengthens a response and omission of reward weakens it, would predict a weaker response with partial reinforcement than with 100% reinforcement in acquisition, performance, and extinction. The facts of post-partial extinction are clearly opposed to this position. By the addition of one further stimulus-response learning-theory concept, it appears that all the facts of behavior under a partial reinforcement schedule can be accounted for. This principle is the common one of generalization, that the more similar the stimulus pattern to the one conditioned, the greater the response strength. These matters are discussed in the following paragraphs.

The first systematic treatment of partial reinforcement appears to have been Skinner's (**62**). In discussing fixed ratio reinforcement, Skinner pointed out that where reinforcement occurs after a pair of responses, the pair should be treated as a unit. He says: "As a rather general statement it may be said that when a reinforcement depends upon the completion of a number of similar acts, the whole group tends to acquire the status of a single response, and the contribution to the reserve tends to be in terms of groups" (p. 300). That is, if two bar-depressions are required to produce reinforcement, and a single reinforcement sets up 10 responses, then 20 single responses (10 pairs) can be expected in extinction after one reward.

According to Skinner, the non-reinforced responses provide a stimulus situation which reinforcement ultimately follows. In other words, in the fixed ratio instance they come to act as stimuli setting the stage for the occurrence of reinforcement. This property is acquired in the same way that external stimuli in a discrimination come to set the occasion for the appearance of reward. The early (non-reinforced) responses in a series such as in the fixed ratio case produce cumulative discriminative stimulation in the presence of which a response is followed by reinforcement. The formation of this discrimination is rarely complete in such instances as the rat's pressing the lever 200 times for a single reinforcement, but the behavior can be acquired and maintained.[8] An aspect of the formation of this discrimination is the spread of secondary reinforcement to the earlier non-reinforced responses in the sequence.

In 1943 Sears (56) set forth a definition of instrumental act that resembles Skinner's earlier formulation in a general way. Sears indicates the difficulty involved in defining "a response" and suggests that when more than one bit of behavior is needed to produce reinforcement, the whole unit should be treated as a response. He specifically points out that more single responses are to be expected in extinction after conditioning where a pair of responses followed by reinforcement is the unit as contrasted with the situation where only single responses are rewarded in training.

Hull (26) does not present a detailed treatment of the problem of partial reinforcement in his book, but in a memorandum to his seminar in the spring of 1941 (25) he discusses what he calls the "Humphreys' paradox." His analysis pertains to Humphreys' (**28**, **31**) findings that resistance to extinction of conditioned eyeblink and GSR is greater after a treatment of semi-randomized 50% reinforcement in conditioning as opposed to 100% reinforcement. Hull's analysis of these results is as follows. For the partial reinforcement group the reaction is conditioned to the after-effects of non-reinforcement. After the first extinction trial the stimulus compound is more like the one conditioned since it now includes the after-effects of non-reinforcement. Presumably the response tendency will be greater to the entire stimulus compound than to any of the parts. From this it is deduced that the effects of preceding extinction trials act to increase response strength above that of the first stimulus presentation of the extinction series for the 50% reinforcement subjects. For groups continuously reinforced in training, the stimulus compound includes only the after-effects of reinforcement, so that extinction trials are much different from conditioning ones. By these means Hull accounts for the initial rise in Humphreys' 50% extinction curves as well as the overall greater resistance to extinction after partial reinforcement as compared with 100%.

---

[8] Questions of the mediation of this kind of activity in terms of verbal behavior are obvious and need not be treated here in detail. With practice the human organism can presumably learn behavior in which exceedingly great blocks or units acquire the properties of single responses. The comptometer operator or typist continues to work his machine for long periods of time involving huge numbers of "single" responses in the absence of obvious external reinforcement. On the basis of verbal cues the following kind of pattern of rewarded (R) and non-rewarded (N) trials, if repeated a number of times, should be learnable: R R N R N N N R R N. Even more complex patterns might well come to serve as response "units." No systematic study seems to be available of the learning of such patterns as response blocks by infra-human organisms.

Miller and Dollard (**43**, p. 39) have pointed out that greater resistance to extinction can be expected when, in conditioning, unrewarded behavior occurs and is ultimately followed by reward. Miller (**42**) has recently elaborated this point of view. The general position relates to several of the approaches already discussed and to the generalization decrement theory to be treated.

An experimental test that reflects several of these viewpoints and particularly Skinner's stand on the response-unit explanation is that of Mowrer and Jones (**47**). Following a suggestion of J. S. Brown, the crux of their position is treatment of a set of non-reinforced responses as a unit of behavior which is strengthened by the terminal reinforcement. Their results indicate fewer response *units* in extinction for partially reinforced groups, although more *single bar-depressions* occurred. Mowrer and Jones' explanation is based on an "effort" hypothesis derived from a previous study concerning variations in the force required to depress the bar (**46**). In the present context they argue that the more effortful a response, the more susceptible it is to extinction. According to this reasoning a unit of four responses requires more effort than a unit of one response.

It should be noted that the fixed ratio reinforcement procedure employed by Mowrer and Jones poses a special discrimination in the free-responding situation in which rapid responding is differentially strengthened by reward. The larger the number of responses required per reward under this schedule, the higher the response strength in conditioning. It is an experimental question whether the effort phase of the response-unit analysis holds for other kinds of partial reinforcement. With periodic reinforcement of bar-pressing, for example, slow responding is differentially rewarded. Yet, greater resistance to extinction may follow.

The other hypothesis invoked to explain these data involves Perin's (52) temporal gradient of reinforcement. Mowrer and Jones (47) cite and support Hull's (26) position that a response more than 30 seconds removed from reward is not strengthened. They indicate that the behavior they studied occurred within this limit. They fail to point out, however, that their fixed ratio procedure was on a very limited scale. Other investigators have employed values 48 times as large in numbers of responses (62) and greater values temporally (36, 37, 60, 62). Also, instances are not accounted for where, for example, the third run down a runway is reinforced and a 90-second inter-trial interval is used. The acceptance by Mowrer and Jones of the 30-second limit on the temporal gradient of reinforcement as a necessary condition for strengthening a response leaves them in the position of failing to account for a considerable body of partial reinforcement data. The generality of this specific phase of the Mowrer-Jones interpretation is exceedingly limited.

It would seem that some mechanism of the response-unit variety is operating. Otherwise, behavior could not be maintained when reinforcements occur only once in nine minutes or every 192 responses. A temporal gradient restricted to strengthening only behavior occurring not more than 30 seconds before reward is, at best, an incomplete account. A temporal gradient may well be one of the factors interacting with several others, but clearly it cannot explain many of the findings that Mowrer and Jones fail to mention.

An important theoretical account of partial reinforcement, somewhat similar to Hull's treatment of Humphreys' data, has recently been proposed and tested by Sheffield (59). Her hypothesis was that the greater resistance to extinction after partial reinforcement is based on the conditioning of the after-effects of non-reinforcement in the stimulus compound during training. In extinction after partial reinforcement, the stimulus situation by way of generalization is more like conditioning than after 100% reinforcement. In extinction, as contrasted with conditioning, the cue pattern is changed greatly for the 100% instance but is reinstated for the partial group. The hypothesis was tested by comparing extinction after distributed and massed training on the assumption that the after-effects of presence or absence of reinforcement in conditioning would be dissipated by spaced training. The results were in striking agreement with the hypothesis: Massed training produced a significant, 30% superiority in resistance to extinction for partial over continuous reinforcement. Distributed training reversed the findings slightly, though not significantly.

Consideration of other partial reinforcement studies in which resistance to extinction was the measure shows almost 100% emphasis on massed training. The only study yielding (an insignificant) superiority of continuous over partial reinforcement in resistance to extinction is Denny's (7), and his data are clearly in line with Sheffield's findings since his training trials were even more widely spaced than her distributed condition.

A study bearing directly on the generalization decrement explanation of the greater resistance to extinction found after partial reinforcement has been performed recently by Sheffield and Temmer (58). These investigators compared the acquisition and extinction of a running response under the escape and avoidance procedures. They point out that extinction practically always involves a change in the cue patterns present in training and particularly so in the omission of a noxious stimulus. Escape training does not allow the conditioning of the consequences of failure to respond; avoidance training maximally connects the response to cue patterns associated with failure to respond since these patterns are the only ones that are reinforced 100% of the time. (This position has recently been discussed by Miller, 42.) The theory predicts greater response strength for the escape procedure in training, but greater resistance to extinction following avoidance conditioning. The data clearly support the theory on both scores.

Although the consequences of the Sheffield-Temmer analysis were tested in the escape-avoidance situation, the approach appears to have wide generality of application to comparisons of partial and 100% reinforcement. It accounts for the partial reinforcement facts in acquisition, performance, and extinction in terms of three central concepts: Presentation of reinforcement strengthens a response, omission of it weakens the response, and a change in the cue pattern is followed by a response decrement by way of generalization. The basic note is that "all partial reinforcement procedures connect the response to cue patterns characteristic of omission of reinforcement, cue patterns which normally contribute—through generalization decrement—to loss of response during extinction after 100% reinforcement. Any procedure which introduces some omission of reinforcement during acquisition should tend to weaken the response during training but prolong extinction of the response" (58).

Further support for this general position is found in a study by Grosslight and Child (18). They point out that the occurrence of a non-rewarded response in a series of rewarded ones permits the response of persisting after failure to be rewarded and strengthened. Such conditioning cannot occur if a failure does not take place in the reward series. The consequences of this position were tested by providing either one or two failures in a series of 10 trials as compared with a control group that received a reward on every trial. The prediction of differential persistence in extinction favoring the failure groups was borne out by the findings.

In a sense, as has been pointed out (47), the situation may be considered in discrimination learning terms. The difference between conditioning and extinction in the 100% instance provides a simply learned discrimination because of the gross dissimilarity between the cues of the two situations. In the partial case, the discrimination is difficult because the cue patterns in conditioning and extinction are similar, i.e., both involve non-reinforcement.

It seems reasonable to assume that a fourth factor is operating in the acquisition and maintenance of partially reinforced responses, namely, derived reinforcement. Non-reinforced responses prior to the reinforced one acquire reinforcing properties since they are ultimately followed by reward. External stimuli support the behavior by way of a similar mechanism. Denny (7) tested the effect of this variable in acquisition by minimizing derived reinforcement in the external environment. When secondary reinforcement was thus reduced, the 100% subjects learned significantly better than the 50% animals; with it operating, the difference between the two groups was considerably less, although still favoring the 100% rats. Denny's results in extinction are confounded by the action of the spacing variable in training. Furthermore, Sheffield (59) has pointed out that for secondary reinforcement to account for the typically greater post-partial extinction curve, it would have to be assumed that secondary reinforcement is stronger than primary reinforcement.

Restricting derived reinforcement in the internal environment poses a more difficult problem. That is, preventing the proprioceptive and kinesthetic stimulation accompanying the response that may well acquire reinforcing properties during training is more troublesome than the parallel restriction in the external environment. Preventing "sensory feedback" from the response by the use of drugs is a possibility.

The overlap between derived reinforcement and the consequences of non-reinforcement in the partial situation is evident. Many of the internal and external cues associated with previous non-reinforcement are present when a reinforced response occurs. Whether the two are separable is an experimental question.

It is noteworthy that the stimulus-response position accounts for certain apparent anomalies, e.g., the greater response strength with less frequent reward under a fixed ratio schedule. This point follows because one basic assumption of this theory is that reward strengthens whatever behavior brings it about. Just as a rat will learn to take the shorter of two paths to food or the one that involves the lesser delay of reward; so it will learn to press faster when more responses are required to produce reward if only because accidentally rapid responding is differentially strengthened.

Expectancy interpretation. The widest use of the principle of expectancy as applied to partial reinforcement situations has been made in a series of investigations by Humphreys (**28, 29, 31, 32**). Hilgard (**23**) has recently amplified the interpretation. In the present context the concept seems to be essentially a common-sense one that may be summarized as follows (**58**). Responses are given because the subject expects (or learns to expect) reinforcement to occur. In extinction after 100% reinforcement the response drops out quickly because the shift from 100% reinforcement to 0% reinforcement leads to a rapid change to expectation of no reinforcement. In post-partial extinction, the subject continues to expect reinforcement to occur, so he continues to respond (**59**).

As Sheffield and Temmer (**58**) have indicated, the expectancy point of view can be made consistent with the facts by means of a few additional assumptions. This interpretation is certainly parsimonious and couched in words meaningful to laymen. Expectancy, as a principle, however, has been widely criticized in the partial reinforcement context (**7, 12, 42, 47, 58, 59, 66**). Three general points emerge: (1) lack of rigorous definition; (2) anthropomorphism; and (3) lack of generality.

As regards definition, Humphreys (**28, 29, 31**) merely notes that expectations are not necessarily verbal in nature. He does not provide a working definition that offers opportunity for predictions and further tests. Sheffield (**59**) has criticized Humphreys most systematically on this score. She points out that the term is not unambiguous in Humphreys' own situations. The greater post-partial resistance to extinction is attributed by him to the difficulty of shifting from an expectation of intermittent reinforcement to one of complete non-reinforcement, as compared with ease in shifting from an expectation of 100% reinforcement to one of 0% reinforcement. Sheffield suggests that the argument can be reversed: After becoming accustomed to partial reinforcement it should be easier to change to an expectation of uniform non-reinforcement, but after 100% reinforcement it should be difficult for the subject to believe that no further reinforcement is forthcoming. In this connection she makes the further point that expectancy as invoked by Humphreys was an *ex post facto* conjecture based on an empirical analysis.

Miller (**42**) has recently suggested that the use of the term expectancy "is circularly based upon the type of behavior it is trying to explain rather than upon certain explicitly stated principles." He goes on to indicate the possibility of rigorous reformulation of the concept for deductive purposes. The further point is made that the term as used in the partial reinforcement context seems to involve circular reasoning: Organisms approach expected rewarding effects and avoid expected noxious effects (**58**).

Humphreys uses the expectancy notion to explain the initial rise he found in post-partial extinction curves. It has been pointed out (**59**) that this statement adds nothing to the conceptualization that reinstating the stimulus pattern conditioned to a response maximizes the probability of the response's occurrence. Furthermore, in other situations the rise is more prominent after *continuous* reinforcement than after partial (**37**, **38**, **62**). In this connection, Detambel (**8**) reports a failure to find the initial rise in a situation similar to Humphreys' (**29**) verbal response study.

The possibility of defining expectancies as implicit verbal responses in human subjects has been pointed out (**47**, **59**). In these terms they would operate like other cues and there would be no need for the additional concept (**59**). The use of such a notion with infra-human organisms is highly questionable; Humphreys himself does not attempt it (**32**), although he notes the possibility of its non-verbal nature (**29**). The difficulties of avoiding the connotations of "expectation" in everyday usage are great; the anthropomorphic implications are obvious (**58**).

Sheffield (**59**) has questioned whether her massed-distributed data can be explained by the expectancy proponents. In her study, both groups learned equally well. Thus, presumably equal expectancies were set up and operated over both the long and short time intervals. Then why should the spaced training wipe out the effects of partial reinforcement in resistance to extinction? Distributed practice should facilitate the acquisition of expectancies as it aids other kinds of acquisition, and these should produce greater extinction differences between partial and continuous groups. But the data indicate the opposite, as predicted by stimulus-response learning concepts.

Mowrer and Jones (**47**) have also suggested a translation of the notion of expectancy into stimulus-response terms. Their point of reference is discrimination learning. The interpretation has been mentioned previously in this paper. It appears to reduce essentially to the generalization-decrement theory, yielding similar predictions and deductions.

The probability theory of Brunswik's (**3**) should be mentioned in the present context. To the extent that his position involves hypotheses and expectations, the same kind of criticism as given previously appears to apply. It may be noted that a simple probability analysis is similar in some respects to one based on strengthening by reinforcement and weakening by extinction (**24**). Other comments on Brunswik's position have been made in Section IV.

VII. Practical Implications

There remains the problem of pulling together the scattered references in this paper to the practical implications and applications of partial reinforcement.[9] Partial reinforcement is certainly more like the

---

[9] Here and elsewhere in this paper we are concerned mainly with conditioning of the instrumental variety. Partial reinforcement has not been studied extensively or system-

Bureau Ednl. Psy. Res.

everyday situation than 100% reinforcement. It is a rare instance in human daily behavior that every response of a specified kind leads to presentation of a reinforcing stimulus. Further, partial reinforcement has the advantage of delaying the effects of satiation. Conditioning can be continued over longer periods with partial as contrasted with 100% reinforcement. The animal trainer uses these techniques. He requires the sequence of tricks to be run through completely or the same act to be repeated several times before feeding the animal. Partial primary reinforcement is not, of course, employed alone in these situations. Maintenance of the behavior leans heavily on derived reinforcement: a pat on the head, a word of praise, etc. The occurrence of this kind of stimulation may be partial or 100%.

The evidence reviewed in the present paper suggests the efficacy of partial reinforcement in everyday life. Learning is not greatly retarded by its use. (It appears, however, to be advantageous to use 100% reinforcement in establishing the desired behavior initially.) The behavior is more than adequately maintained with infrequent reward.[10] Finally, the most striking aspect of the data is the far greater resistance to extinction found after partial reinforcement as compared with continuous. If one desires behavior to be maintained for long periods of time in the absence of externally presented reward, partial reinforcement should be used in training. This generalization appears to hold whether one is interested in having a rat press a bar or in maintaining a child's proper toilet habits or table manners. Implications of the procedure for the practical prediction and control of behavior are obvious.

VIII. Summary and Conclusions

The available literature on partial reinforcement has been reviewed, not merely as a compilation of research facts and findings, but rather as a basis for abstracting the major empirical principles that can be drawn from investigations of this topic. Throughout this article the term "partial reinforcement" has been used to refer to the incidence of reward on less than 100% of the trials or responses. The baseline is 100% or continuous reinforcement.

---

atically in the Pavlovian substitutive situation [with the exception of Brogden's (1) work], although every test trial in which the US is omitted produces a case of noncontinuous reinforcement. Most of the research in this area has involved responses of the skeletal musculature rather than behavior innervated by the autonomic nervous system.

[10] Also, it should be noted that if the reward involved is expensive in time or money, partial reinforcement is the more profitable procedure for maintaining behavior.

The history and development of this research area have been touched on, and the various techniques for presenting partial reinforcement have been treated. Problems in design are mentioned. The main sections of the paper deal with the effects of partial reinforcement on acquisition, maintenance of behavior or performance, and resistance to extinction. A theoretical section that treats the available explanations of partial reinforcement is included.

Empirical generalizations following from the data are as follows:

1. *Acquisition.* Response strength is built up somewhat more rapidly under a schedule of 100% reinforcement than under a partial regimen. Differences in learning, however, are not always large, and with prolonged training the ultimate level of acquisition for partially rewarded subjects may approach that for the 100% ones.
2. *Maintenance.* While the behavior in post-acquisition performance is stable in the partial reinforcement situation, it is usually at a lower level than in the 100% instance. Nevertheless, the differences are not always statistically significant and may well be of no great practical consequence.
3. *Resistance to extinction.* The most striking effects of partial reinforcement are apparent in response strength as measured by resistance to extinction. In almost every experiment, large and significant differences in extinction favoring the groups partially reinforced in conditioning over the 100% ones were found. The practical implication of this principle for maintaining behavior is obvious: Administer the reinforcing stimulus in conditioning according to a partial schedule, and the behavior will be maintained for long periods in the absence of external support from primary reward.

Stimulus-response learning theory concepts and those stemming from an expectancy point of view are discussed as explanations of partial reinforcement data. It appears that S-R principles provide a framework accounting for the phenomena in this area.

BIBLIOGRAPHY

1. Brogden, W. J. The effect of frequency of reinforcement upon the level of conditioning. *J. exp. Psychol.*, 1939, **24**, 419–431.
2. Brown, J. S. A note on a temporal gradient of reinforcement. *J. exp. Psychol.*, 1939, **25**, 221–227.
3. Brunswik, E. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, **25**, 175–197.
4. Brunswik, E. Organismic achievement and environmental probability. *Psychol. Rev.*, 1943, **50**, 255–272.
5. Calvin, J. S. Learning of probable occurrences. *Amer. Psychologist*, 1948, **3**, 236. Abstract (manuscript seen).
6. Cole, L. E. A comparison of the factors of practice and knowledge of experimental procedure in conditioning the eyelid response of human subjects. *J. gen. Psychol.*, 1939, **20**, 349–373.
7. Denny, M. R. The role of secondary reinforcement in a partial reinforcement learning situation. *J. exp. Psychol.*, 1946, **36**, 373–389.
8. Detambel, M. H. A reanalysis of

Humphreys' verbal expectations. Unpublished M.A. thesis, Indiana Univ., 1950.

9. DOUGLAS, HELEN B., & BROWN, JANET L. The effect of heterogeneous reinforcement procedures on the extinction curve of the white rat. Unpublished study, Mt. Holyoke Coll., 1949.
10. EHRENFREUND, D. The effect of partial reinforcement on the strength of secondary reinforcing cues. *Amer. Psychologist*, 1948, **3**, 241. Abstract (title only—manuscript seen).
11. ESTES, W. K. An experimental study of punishment. *Psychol. Monogr.*, 1944, **57**, No. 3, 1–40 (whole No. 263).
12. FINGER, F. W. The effect of varying conditions of reinforcement upon a simple running response. *J. exp. Psychol.*, 1942, **30**, 53–68.
13. FINGER, F. W. Retention and subsequent extinction of a simple running response following varying conditions of reinforcement. *J. exp. Psychol.*, 1942, **31**, 120–133.
14. FRANZ, ELAINE G., FURNIVALL, NANCY A., & JORDAN, JEANNE M. A comparison of the effects of regular and aperiodic reinforcement on the conditioning of a new operant with secondary reinforcement. Unpublished study, Mt. Holyoke Coll., 1949.
15. GILINSKY, ALBERTA S., & STEWART, JEANNIE C. "Extinction" of a success aspiration following three conditions of reinforcement. *Amer. Psychologist*, 1949, **4**, 222–223. Abstract.
16. GOTTSDANKER, R. M. An experimental study of fixation of response by college students in a multiple-choice situation. *J. exp. Psychol.*, 1939, **25**, 431–444.
17. GRANT, D. A., & HAKE, H. W. Acquisition and extinction of the Humphreys' verbal response with differing percentages of "reinforcement." *Amer. Psychologist*, 1949, **4**, 226. Abstract.
18. GROSSLIGHT, J. H., & CHILD, I. L. Persistence as a function of previous experience of failure followed by success. *Amer. J. Psychol.*, 1947, **60**, 378–387.
19. HALL, J. F., & WICKENS, D. D. Secondary reinforcement as a function of the number of differential primary reinforcements. *Amer. Psychologist*, 1949, **4**, 226. Abstract.
20. HEATHERS, G. L. The avoidance of repetition of a maze reaction in the rat as a function of the time interval between trials. *J. Psychol.*, 1940, **10**, 359–380.
21. HERON, W. T. The effects of a differential rate of reinforcement of responses to two levers. *J. comp. Psychol.*, 1942, **33**, 87–96.
22. HERON, W. T., & SKINNER, B. F. Changes in hunger during starvation. *Psychol. Rec.*, 1937, **1**, 51–60.
23. HILGARD, E. R. *Theories of learning.* New York: Appleton-Century-Crofts, 1948.
24. HILGARD, E. R., & MARQUIS, D. G. *Conditioning and learning.* New York: Appleton-Century, 1940.
25. HULL, C. L. *Psychology seminar memoranda.* 1940–1941. Bound mimeographed copies on file in the libraries of the Univ. of Iowa, Univ. of North Carolina, and Yale Univ.
26. HULL, C. L. *Principles of behavior.* New York: Appleton-Century, 1943.
27. HULL, C. L., & SPENCE, K. W. "Correction" vs. "non-correction" method of trial-and-error learning in rats. *J. comp. Psychol.*, 1938, **25**, 127–145.
28. HUMPHREYS, L. G. The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25**, 141–158.
29. HUMPHREYS, L. G. Acquisition and extinction of verbal expectations in a situation analogous to condition-

ing. *J. exp. Psychol.*, 1939, **25**, 294–301.

30. Humphreys, L. G. Generalization as a function of method of reinforcement. *J. exp. Psychol.*, 1939, **25**, 361–372.
31. Humphreys, L. G. Extinction of conditioned psycho-galvanic responses following two conditions of reinforcement. *J. exp. Psychol.*, 1940, **27**, 71–75.
32. Humphreys, L. G. The strength of a Thorndikian response as a function of the number of practice trials. *J. comp. Psychol.*, 1943, **35**, 101–110.
33. Hunter, W. S. The temporal maze and kinaesthetic sensory processes in the white rat. *Psychobiol.*, 1920, **2**, 1–18.
34. Jenkins, W. O. Drive and resistance to extinction. (In prep.)
35. Jenkins, W. O., & Clayton, Frances L. Rate of responding and amount of reinforcement. *J. comp. physiol. Psychol.*, 1949, **42**, 174–181.
36. Jenkins, W. O., McFann, H., & Clayton, Frances L. A methodological study of extinction following aperiodic and continuous reinforcement. *J. comp. physiol. Psychol.*, 1951, **44**, (in press).
37. Jenkins, W. O., & Rigby, Marilyn K. Partial (periodic) vs. continuous reinforcement in resistance to extinction. *J. comp. physiol. Psychol.*, 1950, **43**, (in press).
38. Keller, F. S. The effect of sequence of continuous and periodic reinforcement upon the 'reflex reserve.' *J. exp. Psychol.*, 1940, **27**, 559–565.
39. Lawrence, D. H., & Miller, N. E. A positive relationship between reinforcement and resistance to extinction produced by removing a source of confusion from a technique that had produced opposite results. *J. exp. Psychol.*, 1947, **37**, 494–509.
40. Lewin, K. *A dynamic theory of personality* (trans. by D. K. Adams & K. E. Zener). New York: McGraw-Hill, 1935.
41. Lloyd, V. V. Resistance to extinction after periodic and aperiodic reinforcement. Unpublished Ph.D. thesis, Columbia Univ., 1950.
42. Miller, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.) *Handbook of experimental psychology.* (In prep.)
43. Miller, N. E., & Dollard, J. *Social learning and imitation.* New Haven: Yale Univ. Press, 1941.
44. Mosteller, F. Personal communication, 1949.
45. Mosteller, F., & Nogee, P. Measurement of utility in a gambling situation. (In prep.)
46. Mowrer, O. H., & Jones, Helen M. Extinction and behavior variability as functions of effortfulness of task. *J. exp. Psychol.*, 1943, **33**, 369–386.
47. Mowrer, O. H., & Jones, Helen M. Habit strength as a function of the pattern of reinforcement. *J. exp. Psychol.*, 1945, **35**, 293–311.
48. Mueller, C. G. Numerical transformations in the analysis of experimental data. *Psychol. Bull.*, 1949, **46**, 198–223.
49. Murphy, M. J. *The differentiation of a response; an exploratory study with human subjects.* Abridgement of a Ph.D. dissertation. New York: New York Univ., 1943.
50. Pavlov, I. P. *Conditioned reflexes* (trans. by G. V. Anrep). London: Oxford Univ. Press, 1927.
51. Perin, C. T. Behavior potentiality as a joint function of the amount of training and the degree of hunger at the time of extinction. *J. exp. Psychol.*, 1942, **30**, 93–113.
52. Perin, C. T. The effect of delayed reinforcement upon the differentiation of bar responses in white rats. *J. exp. Psychol.*, 1943, **32**, 95–109.
53. Platonov, K. I. The formation of associated motor reflexes to simultaneous auditory and visual stimuli in human beings. Unpublished

thesis, St. Petersburg Univ., Russia, 1912. Cited by Razran.

54. Razran, G. H. S. Conditioned withdrawal responses with shock as the conditioning stimulus in adult human subjects. *Psychol. Bull.*, 1934, **31**, 111–143.
55. Saltzman, I. J. Maze learning in the absence of primary reinforcement: a study of secondary reinforcement. *J. comp. physiol. Psychol.*, 1949, **42**, 161–173.
56. Sears, R. R. Survey of objective studies of psychoanalytic concepts. *Soc. Sci. Res. Coun. Bull.*, No. 51, 1943.
57. Sheffield, F. D. Avoidance training and the contiguity principle. *J. comp. physiol. Psychol.*, 1948, **41**, 165–177.
58. Sheffield, F. D., & Temmer, Helena W. Relative resistance to extinction of escape training and avoidance training. *J. exp. Psychol.*, 1950, **40**, (in press).
59. Sheffield, Virginia F. Extinction as a function of partial reinforcement and distribution of practice. *J. exp. Psychol.*, 1949, **39**, 511–525.
60. Skinner, B. F. The rate of establishment of a discrimination. *J. gen. Psychol.*, 1933, **9**, 302–350.
61. Skinner, B. F. The extinction ratio and its modification by a temporal discrimination. *Psychol. Bull.*, 1936, **33**, 784. Abstract.
62. Skinner, B. F. *The behavior of organisms.* New York: Appleton-Century, 1938.
63. Skinner, B. F. Personal communication, 1949.
64. Skinner, B. F., & Heron, W. T. Effects of caffeine and benzedrine upon conditioning and extinction. *Psychol. Rec.*, 1937, **1**, 340–346.
65. Spence, K. W. The differential response in animals to stimuli varying within a single dimension. *Psychol. Rev.*, 1937, **44**, 430–444.
66. Stanley, J. C., Jr. The differential effects of partial and continuous reward upon the acquisition and elimination of a running response in a two-choice situation. Unpublished Ed.D. thesis, Harvard Univ., 1950.
67. Tolman, E. C. The determiners of behavior at a choice point. *Psychol. Rev.*, 1938, **45**, 1–41.
68. Tolman, E. C., & Brunswik, E. The organism and the causal texture of the environment. *Psychol. Rev.*, 1935, **42**, 43–77.
69. Wentink, Elaine A. The effects of certain drugs and hormones upon conditioning. *J. exp. Psychol.*, 1938, **22**, 150–163.
70. Williams, S. B. Resistance to extinction as a function of the number of reinforcements. *J. exp. Psychol.*, 1938, **23**, 506–522.
71. Wolfle, Helen M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, **7**, 80–103.

Received November 15, 1949.

THE FACTORIAL ANALYSIS OF ANIMAL BEHAVIOR[1]

JOSEPH R. ROYCE[2]
University of Chicago

INTRODUCTION

The primary purpose of the factorial approach in the study of behavior is to determine the basic parameters which are common to the interrelated variables of a given domain. Obviously then, it is a matter of simplification, an application of the law of parsimony to complex data whose interrelationships are difficult to comprehend. More specifically, in the case of individual differences, a limited number of abilities or factors are isolated as the determiners of behavior in a given domain. Recalling that a variable may be expressed in standard form, Z_j, (where j equals any variable) and denoting common factors by the notation F_j and their coefficients by a_j, we have the following representation of the preceding statement for three of the total number of variables:[3]

$$Z_1 = a_{11}F_1 + a_{12}F_2 + a_{13}F_3 + \cdots + a_{1m}F_m. \qquad [1]$$

$$Z_2 = a_{21}F_1 + a_{22}F_2 + a_{23}F_3 + \cdots + a_{2m}F_m. \qquad [2]$$

$$Z_3 = a_{31}F_1 + a_{32}F_2 + a_{33}F_3 + \cdots + a_{3m}F_m. \qquad [3]$$

To express the variance of a particular variable, which, it will be recalled, is equal to σ^2, square both sides of [1], sum over the n values of the respective variables, and divide by N, thus obtaining for [1]:

$$\frac{\Sigma Z_1^2}{N} = a_{11}^2\frac{\Sigma F_1^2}{N} + a_{12}^2\frac{\Sigma F_2^2}{N} + a_{13}^2\frac{\Sigma F_3^2}{N} + \cdots + a_{1m}^2\frac{\Sigma F_n^2}{N}. \qquad [4]$$

Now, since the variance of a variable in standard form is equal to unity, equation [4] may be written:

$$\sigma_{Z_1}^2 = 1 = a_{11}^2 + a_{12}^2 + a_{13}^2 + \cdots + a_{1m}^2. \qquad [5]$$

Thus, the terms on the right (a_{11}^2, etc.) represent the portions of unit variance which are accounted for by each of the respective common factors (F_1, F_2, F_3).

From this it is obvious that if a given variable were a "pure" measure of a particular factor, say F_1, the a_{12}^2 and a_{13}^2 etc. values would be equal to zero, and the variance of the variable would be entirely accounted for by one common factor (F_1). In essence, the determination of the factor coefficients (a_1, a_2, $a_3 \cdots + a_m$, i.e., factor loadings) is the basic problem of factor analysis. On the human level, Thurstone has isolated several determiners of mental ability (50) to account for individual differences in this domain. In recent years, similar attempts have been made in certain areas of animal behavior, although paradoxically, no studies have been made of the higher mental processes in animals. This is undoubtedly at least partially due to the lack of evidence of "higher" mental activity among lower animals until relatively recent years. It is also due to the fact that factor methodology did not attain sufficient dignity until about 1935. In any event, factorial analyses of animal behavior are being carried out, and although the factors thus far isolated need further verification, the writer feels justified in

---

[1] The writer would like to thank Professor K. S. Lashley and Professor L. L. Thurstone for critical reading of this paper in manuscript.

[2] Now at Drake University, Des Moines, Iowa.

[3] For the sake of simplicity this treatment is concerned only with orthogonal common factors; the uniqueness of each variable is omitted.

TABLE I

INTERCORRELATIONS BETWEEN EARLY MEASURES OF LEARNING IN ANIMALS

Experimenter	*Date*	*Animals*	*Tests Used*	*Correlation*
Bagg (4)	1920	90 mice	Simple maze vs. multiple choice	.11
Heron (25)	1922	22 rats	Circular maze vs. problem box	.01 to .09
Hunter and Randolph (32)	1924	23 to 37 rats	T maze, sawdust box, straightaway	−.16 to .18
Tolman & Davis (55)	1924	13 rats	Rectangular maze vs. Carr maze	.16 to .66
Liggett (35)	1925	48 chicks	T maze straightaway	−.64 to .19
Williams (68)	1929	25 rats	Visual discrimination vs. multiple T-maze	.16 and .08
Miles (38)	1930	38 rats	Errors, elevated vs. alley multiple T maze	.50
Tryon (58)	1931	141 rats	17 units vs. 20 unit multiple T maze	.77

presenting the work done and in pointing out several reasons why such studies are of crucial importance.

EARLY CORRELATION STUDIES

The earliest attempts to reveal interrelationships of animal performance were carried out between two measures of a specific learning task. Partly because of unreliable and invalid measures, and partly because of inadequate sampling, these early investigations revealed near zero intercorrelations. It was not until the studies of Miles (**38**) and Tryon (**58**) that appreciable relationships were shown to exist. A summary of these early studies is reproduced in Table I.

TABLE II

INTERCORRELATIONS OF LEARNING MEASURES OF THE RAT (9)

	Multiple T-Maze	*Elevated U-Maze*	*Elevated T-Maze*	*Stone Light Discrimination*
Multiple T-Maze		.63	.53	−.10
Elevated U-Maze			.61	.01
Elevated T-Maze				−.02

TABLE III

INTERCORRELATIONS OF LEARNING MEASURES OF THE RAT (9)

	Triple Platform Problem Box	*T-Maze*	*Light Discrimination*
Problem Box		−.06	−.05
T-Maze			.01

In the early thirties several multiple variable studies were made, primarily by the Stanford group of animal psychologists. These exploratory investigations were not very extensive in scope, although they were all relevant to learning performance. The essence of these studies is presented in Tables II, III and IV. Although these data show that intercorrelations between maze performances are all positive and significant, and that intercorrelations between other similar measures of learning are positive, it is evident that interrelationships between distinctly different measures of learning tend toward zero. Thus, there is no evidence of what might be called a general learning factor. Comparable findings have been reported by Woodrow (**70, 71**) who was unable to find a general learning factor in human subjects. Furthermore, a study by Campbell (8) indicates that this same specificity of animal behavior is evident in a comparison of rat behavior in "learning" situations and "reasoning" situations. Campbell's results concerning the relationships between maze performance and reasoning performance are shown in Table V.

TABLE IV

INTERCORRELATIONS OF LEARNING MEASURES OF THE RAT (53)

	Warden U-Maze	*U-Maze Reversed*	*Miles Elevated Maze*	*Miles Maze Reversed*	*Multiple Light Discr.*	*Discr. Reversed*
Warden U-Maze		.55	.47	.33	.03	−.05
U-Maze Reversed			.52	.41	−.03	−.19
Miles Elevated Maze				.51	−.01	−.12
Miles Maze Reversed					.09	.00
Multiple Light Discr.						.52

TABLE V

RELATIONSHIP BETWEEN LEARNING PERFORMANCE AND PERFORMANCE ON THE MAIER REASONING TEST (8)

	Stone Multiple T-Maze	*Warden U-Maze*	*Maier Reasoning*
Stone Multiple T-Maze		.55	−.37 (−.29)
Warden U-Maze			−.22 (−.20)

FACTORIAL STUDIES OF LEARNING

The first frontal attack on the underlying structure of learning in animals was made by Dunlap (11) in 1933. Since most of the experimental work on individual differences has been performed on rats, and this original study was done with chicks, it will not be treated in detail. Briefly, Dunlap submitted 119 chicks to a battery of ten tests, mostly learning tasks. The obtained correlation matrix (average r around .20) was subjected to tetrad analysis to determine whether the underlying traits could be analyzed into a single general factor plus specifics. Such a pattern was found inadequate and the non-existence of a general learning ability was again confirmed, this time in chicks.

McCulloch (**37**) was interested in the same problem in rats. He obtained results of the performance of 85 rats on eight learning tasks, all of which involved the "apprehension of relations."[4] The battery consisted of six discrimination tests, one 27 unit elevated maze, and a two-pedai problem box, plus an irrelevant test of motor dexterity. As can be seen in Table VI, except for one *r* of .44 (between two light discrimination tasks), the average *r* is around .016. Obviously, the magnitude of the intercorrelations precludes the existence of a "*g*" factor to be extracted from these data.

TABLE VI

INTERCORRELATIONS OF MEASURES OF LEARNING OF THE RAT (**37**)

	I	II	III	IV	V	VI	VII	VIII	IX
I									
II	44								
III	11	09							
IV	03	11	20						
V	04	−05	11	−30					
VI	−16	−09	−16	−18	01				
VII	−04	−24	13	15	11	08			
VIII	−04	−04	−01	08	08	−11	−27		
IX	−14	−22	11	−05	06	19	15	−18	

Robert Thorndike made the first comprehensive investigation of rat behavior which involved an application of the Thurstone centroid method of factor analysis. His purpose was to cover as wide a variety of learning problems as possible. For this rather extensive task he used seven experimental set-ups from which he derived 32 scores (average inter *r* of .18). The apparatus were as follows: activity wheel, Warner-Warden maze (two patterns), elevated T maze (two patterns), Jenkins circular problem box, latch problem box, Warner conditioned response test (two stimuli), and the Columbia obstruction apparatus. Several analyses were made, one using the Hotelling solution, one the Thurstone solution. Also, analyses were made on three different matrices, one of 13 variables, one of 20 variables, and one of 32 variables. Although the loadings differed in detail, the basic patterns were the same. It was concluded that three factors accounted for the total variance to be analyzed. These Thorndike (**49**) labeled as follows: docility (common to all

[4] What Spearman in his two-factor theory has referred to as eduction.

tests), transfer (distinguishing early from later tests), and a conditioned response factor (specific to conditioned response scores).

In the opinion of several investigators, Thorndike's results were not regarded as particularly meaningful. It was pointed out that although Thorndike used the centroid method, he did not rotate his obtained factor loadings to test for simple structure, and hence, to reach a more meaningful solution. As a consequence, Thorndike's data were re-analyzed several years later by Van Steenberg (**63**). In methodology, the essential difference in this study is the rotation of axes to a more meaningful configuration. In this analysis, ten factors were isolated, only five of which were interpreted. These five, which are primarily relevant to learning, have been labeled as follows: visual (elevated mazes), adaptability to test situations, speed, kinesthetic ability, and visual insight (maze Gestalt).

These investigations were followed by a similar study by Vaughn (**64**) in 1937. Judging by the battery of tests employed, the purpose of the investigator was to explore the entire rat behavior domain by use of tests of known reliability and validity. The ten pieces of apparatus employed were: wildness tunnel, activity cage, straightaway, simple maze, multiple T-maze, rectangular maze, problem box, Maier "reasoning" test, and a perseverance box. In the factorial analysis of the 34 variables obtained, the investigator in this experiment employed the centroid method with rotation to test for simple structure. Of the eight factors isolated, four were sufficiently clear to be interpreted. The labels were given as follows: speed, wildness-timidity (docility), insight learning, and transfer. Again, the essential nature of the factors obtained, with the possible exception of wildness-timidity, refer to the organization of learning behavior.

A comparison of the several studies thus far reported indicates several trends which will perhaps be made clear by the summary presented in Table VII. These data, coupled with the indices of earlier correlation studies lead to several fairly clear relations. The most obvious conclusion is that the assumption of a unitary learning factor is untenable. This is most clearly indicated in the factorial studies of Dunlap (**11**) and McCulloch (**37**); particularly in McCulloch's investigation which was specifically directed at Spearman's original postulation of "*g*." An alternative explanation is that several factors are involved in the organization of learning. The subsequent studies of R. L. Thorndike (**49**), Van Steenberg (**63**), and Vaughn (**64**) attest to this interpretation, although the exact nature of such group factors is not certain at present.

The data presented thus far have all been concerned with learning performance regarded as a static affair. Wherry (**65, 66, 67**) has initiated a unique series of studies which gets at the dynamics of learning by clever usage of the factorial methods. In the first place, Wherry claims that a mere analysis of composite scores of behavior for an entire learning situation will not give a complete picture of what goes on. He postulates the idea that many simultaneous learning processes are taking place at the same time within a given task. Thus, each of these may have learning mechanisms which differ from one another. For example, it may be possible that a certain factor which accounts for performance for a few trials when it is dominant may be relatively unimportant in a later part of the learning situation when another factor dominates. It is this very hypothesis which his studies were carried out to test. In order to get at the entire learning process, his analyses were applied to correlation coefficients obtained between the separate trials or stages of the learning period (i.e. period 1 correlated with period 2, period 3; period 2 correlated with period 3, period 4, etc.). Thus, various alleys were considered as individuals, and the number of entrances into each alley was considered as the score for that day, trial, or period. The intercorrelation matrix obtained in this fashion was then tested by the centroid method and rotated for simple structure. This procedure was carried out on five sets of experimental learning data which had been gathered by other investigators. The data include 2 learning experiments described by Yoshioka (**72, 73**)—pattern discrimination and distance discrimination—Hunter's (**30**) simple and complex

TABLE VII

SUMMARY OF FACTORS ISOLATED IN RAT LEARNING DOMAIN

No. of Factors Isolated	*Identification of Factors*	*Investigator*	*Year*
0	no "g"	McCulloch (**37**)	1935
3	Docility Transfer CR factor	R. Thorndike (**49**)	1935
5	Visual Adaptability Speed Kinesthetic Visual insight	Van Steenberg (**63**)	1937
4	Speed Docility Insight learning Transfer	Vaughn (**64**)	1937

maze performance, Ruch's (47) linear mazes, and Tryon's (58) 17 unit multiple-T maze for both bright and dull rats. In my estimation the results are quite striking. Furthermore, they are consistent! Exactly the same pattern was found regardless of the particular learning problem confronting the animals. Samples of these typical results are reproduced in Figs. 1 and 2. At this stage of the knowledge of such factors it would appear that the specific description of the factors obtained would depend upon the nature of the problem confronting the animal. In any event, Wherry (65) has offered a general description which adequately accounts for the factors obtained regardless of its specific setting. There are inferences that the initial factor corresponds to trial and error running, that the rising and waning factor involves retracing and uncertain behavior, and that the last factor represents the insightful learning of the correct response. However, Wherry (pp. 271–272) cautiously makes only the following conclusions: (1) These preliminary studies have shown that factors present in learning situations are not

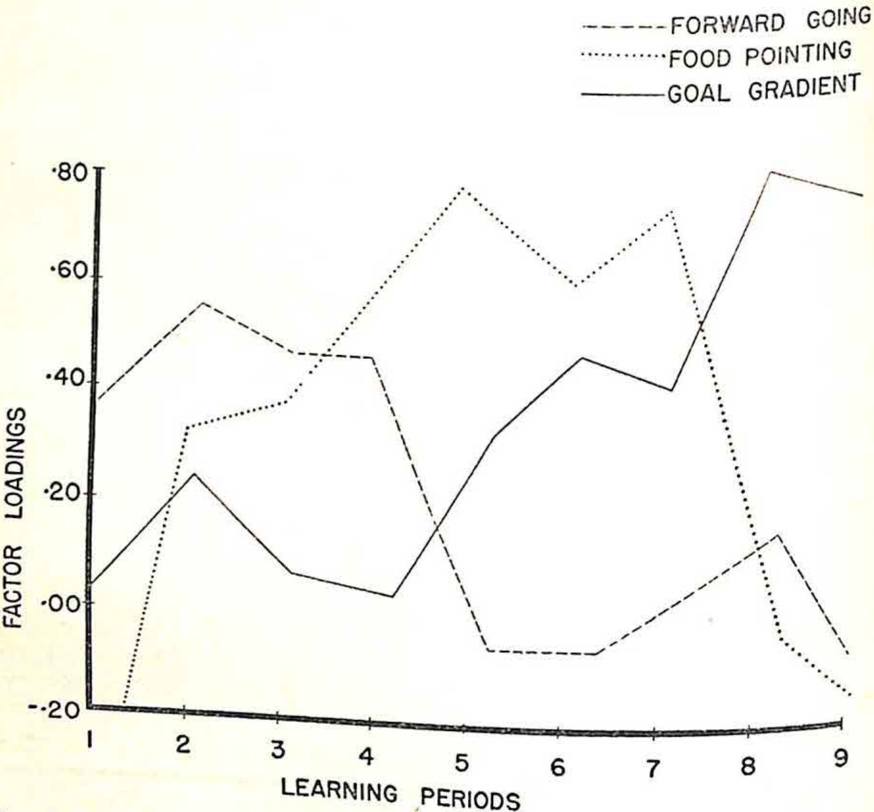


Fig. 1. Factors obtained from Yoshioka Discrimination Data (65, p. 267).

static affairs, acting uniformly throughout the learning of the maze. (2) Furthermore, it is possible by means of proper factorial analysis to separate the effects of these various factors, thus enabling us to see the course of their rise and wane as well as giving us a less confusing picture of their importance.

In another paper, Wherry (67) analyzes the important genetic studies of individual differences in rat maze behavior which Tryon has been conducting over a period of about twenty years. Wherry offers a more parsimonious description of the ten components suggested by Tryon (60) and then goes on to an analysis of the differential effects of the obtained factors in the learning situation. The results obtained, again, are consistent with results previously obtained by Wherry. This fact is immediately obvious by inspection of Figs. 3 and 4. In addition to pointing up the consistency of Wherry's results, these data are of interest on several other counts. In the first place, it will be recalled that the animals used in Tryon's original studies were selectively bred

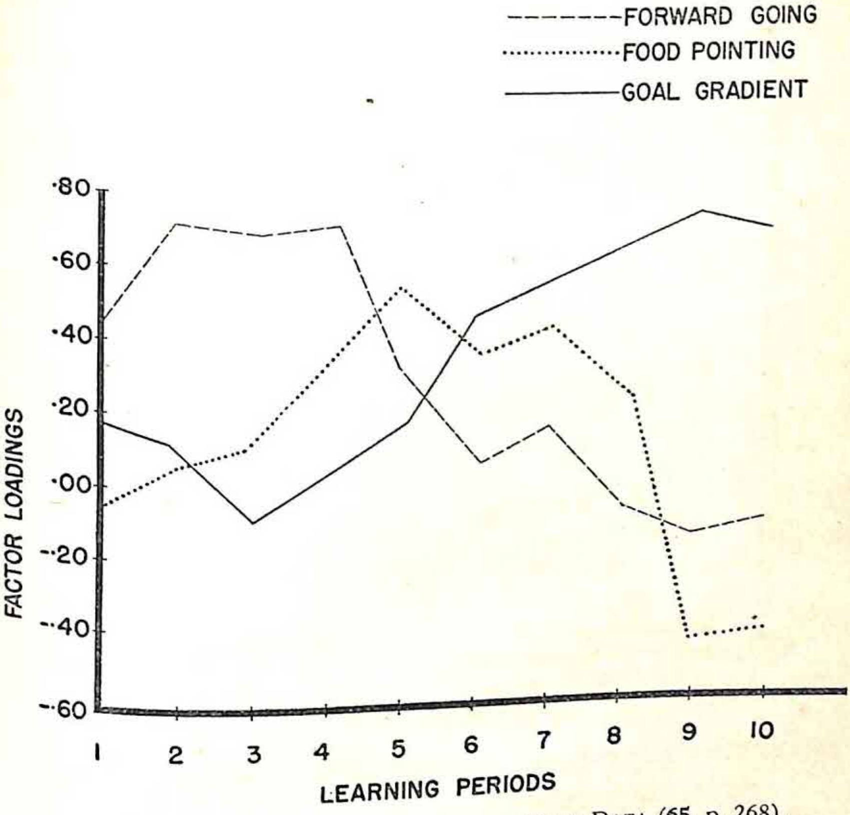


FIG. 2. FACTORS OBTAINED FROM HUNTER MAZE DATA (65, p. 268).

ried out by Hall and his associates. Work in this area has obviously been slow in getting started because of the great difficulty in establishing a useful index of emotionality. However, in 1934, Hall (**13**) reported an index of emotionality which has apparently withstood the test of time and criticism. Because of its importance with respect to later studies, a brief treatment of this investigation will be given. Hall begins by explaining his preference for a concept of emotionality rather than emotion.

TABLE VIII

NUMBER AND PERCENTAGE OF RATS DEFECATING AND EATING FOR EACH TRIAL (**13**, p. 388)

Days	*Defecating*		*Eating*	
	No.	*Percent*	*No.*	*Percent*
1	18	69	2	08
2	18	69	2	08
3	17	66	5	19
4	10	38	7	27
5	13	50	8	31
6	10	38	8	31
7	7	27	13	50
8	6	23	16	61
9	4	15	18	69
10	4	15	20	77
11	3	11	18	69
12	3	11	20	77
13	2	08	22	84
14	2	08	24	92

> This state consists of a group of organic, experiential and expressive reactions and denotes a general upset or excited condition of the animal. Emotionality can be thought of as a trait since animals and men differ in the intensity of emotional reactions displayed. The reader is warned against interpreting emotionality as a thing or faculty. It is merely a convenient concept for describing a complex of factors. Emotionality is preferable to the term emotion since the latter implies for most psychologists that a differentiation of emotions can be made. The current point of view is that such differentiation is extremely speculative (**13**, p. 385).

With this conception in mind, Hall proceeds to establish his measure for individual differences in emotionality among rats. He placed 31 rats in a strange "open field" for three minutes a day for twenty days.

Lettuce was scattered within the enclosure, and the rats were appropriately motivated. Defecation, urination, and eating were recorded. The following highly interesting results were obtained (Table VIII). It is obvious from this table that there is a significant relationship between emotional defecation and eating. Initially the percentage of animals defecating is high (69%) and the percentage of animals eating is low (8%). As the percentage of animals defecating decreases, the percentage of animals eating increases. Expressed numerically, Hall reports a correlation of .82 between "days defecating" vs. "days not eating." A similar relationship obtains between "days urinating" vs. "days not eating" (*r* equals .70). The raison d'être for the rats' emotionality in the field situation is given as the unfamiliarity or strangeness to the animal. The reasoning behind the establishment of eating as the criterion of emotionality is succinctly stated by Hall as follows:

> The problem of finding another measure of individual differences in emotionality (criterion) was not easy, and was finally solved by the following reasoning. It is a commonplace that an upset condition hinders the adjustment of the animal. Emotionality may act as a barrier to the satisfaction of needs. Hence it was reasoned that if a rat which would eat under normal living cage conditions immediately upon the presentation of food was placed in a strange enclosure, where food was readily obtainable, its failure to eat could be said to be due to the emotionality of the animal. The number of trials it required before the food was eaten could be taken as a measure of the rat's emotionality (**13**, pp. 389–90).

Hall concludes that emotional defecation and urination are valid measures of individual differences in emotion. It should be noted that one of the stronger points in favor of such a criterion of emotionality lies in the fact that such defecation and urination are under the control of the autonomic nervous system.

In a subsequent series of studies (**14**, **15**, **16**, **17**, **18**) from 1934 to 1941 Hall investigated interrelationships between emotionality and need, ambulatory activity, stereotyping of behavior, persistence, and behavior derangements. His primary conclusions are that emotional disturbance tends to be reduced when the rat is impelled by a strong need, emotional rats are less active than relatively non-emotional rats, the emotional rat is less stereotyped (more variable) in his behavior, emotional rats are less susceptible to audiogenic attacks than the non-emotional rats, and that persistence and emotionality are not related temperamental traits. In a later study (**23**), in which genetically derived strains of emotional vs. non-emotional strains were selectively bred, Hall provides evidence for the inheritance of aggressiveness in the non-emotional strain and timidity in the emotional strain.

These studies culminated in an intercorrelational study by Billingslea (5) in 1941. In this study the investigator studied the relationship between emotionality and various other allied indices. The measures included Hall's open field situation, an activity wheel, a modified Vaughn paper barrier (persistence), coordination of paws and mouth in an enclosed food dish, aggressiveness, wildness-timidity, and the Martin-Hall standard situation for measuring abnormal behavior. The subjects were 40 adult female rats, 20 from Hall's emotional strain and 20 from the non-emotional strain. In a later paper (7), Billingslea subjected the obtained intercorrelation matrix to a factorial analysis. Three factors were postulated, and identified as emotionality, freezing, and timidity. The factor loadings are reproduced in Table IX. The largest factor loading for each factor is marked with an asterisk to indicate the essence of the factors identified. It is obvious that similarly high loadings vary in the same direction. Although much remains to be done, the results are encouragingly clear and meaningful.

TABLE IX

FACTOR COEFFICIENTS OF EMOTIONALITY FACTORS (6, p. 209)

Variables	*Emotionality*	*Freezing*	*Timidity*
Emotionality	.616*	.348	−.227
Activity	.449	−.271	−.358
Paw-mouth	−.017	−.161	−.030
Fighting	−.008	−.040	−.619*
Probl. Solving (persistence)	−.082	.652*	.433
Aggressiveness (air stream)	−.389	−.502	.106
Wire T. (Time in start cage)	.450	.023	.279
Wire tunnel (time in tunnel)	.403	.141	.398
Pipe (S. C. time)	−.026	.528	.553
Pipe Tunnel (time)	.023	.407	.382
Behavior Disturbance	−.003	−.535	.033

STUDIES OF MOTIVATION

With the exception of one study by Rethlingshafer (43) all large scale intercorrelational studies of motivation and drives reported thus far have been conducted by Anderson (1, 2, 3). The most thorough of the three interrelated studies involves 47 measures obtained from 14 experimental set-ups. Without specifying the apparatus employed, the measures obtained may be summarized as follows: exploration, 7; hunger, 14; thirst, 10; sex, 7; learning, 6; and one measure each of body weight, activity, and emotion. Although Anderson had originally intended to carry through with a factorial analysis of the obtained intercorrelation matrix, he did not perform such an analysis. His attitude on the necessity for such a procedure is perhaps indicated in his first paper wherein he states that "methods of factor analysis have been developed in an effort to determine just what factors may be involved in any given test, but at the present stage of development of the methods (1934) the analyses are often more confusing than illuminating because the factors fail to make good psychological sense" (**1**, p. 107). Although there was some justification for this point of view when factor theory was in its early stages, there are indications that these same data could well be analyzed factorially by current methods in an effort to obtain greater clarity of the underlying relationships. This is particularly true in areas where Anderson's interpretations are admittedly speculative. One outstanding instance of this is apparent in the following conclusion wherein he states: "There is a fairly general tendency for sex tests, learning measures and the emotional defecation test to yield significant or almost significant intercorrelations. There is thus some evidence of the operation of a general factor or factors within these groups of tests" (**2**, p. 113). In any event, certain generalizations were apparent from the data; they may be briefly summarized as follows:

1. Exploratory tests correlated significantly with other exploratory tests.
2. Sex tests correlated significantly with other sex tests.
3. Thirst tests did not correlate significantly with other thirst tests.
4. Hunger tests did not correlate significantly with other hunger tests.
5. Learning measures correlated with other learning measures.
6. No general drive factor is in evidence between measures of different drives.
7. Sex tests, learning measures, and emotional defecation are interrelated.
8. Learning measures show little tendency to correlate with motivational measures.

Further generalizations which need to be confirmed by factorial analysis are: "The data obtained indicate that the performance on any given test may be a complex function of a number of factors, many of which are probably quite independent of the particular drive involved" (**2**, p. 98). And "one is not justified in assuming that any given drive is necessarily a unitary factor operating to produce significant and consistent intercorrelations among different tests of the specified drive" (**2**, p. 107). The only safe generalization that can be made at this point is that the different drives are independent and are not related to learning ability.

One other multiple variable investigation in the motivation area has been reported by Rethlingshafer (43). In essence, this study is an application of the Wherry procedure of applying factor analysis to periodic learning. In this instance the investigator was concerned with differences in the course of learning under varying motivating conditions in the learning of a visual discrimination habit. The original experimental data were gathered by Muenzinger (39) who used two sets of motivating conditions: (a) producers of fast learning: shock-wrong, shock-right, shock-after-choice, jump-after-choice, and delay-at-the-point-of-choice;

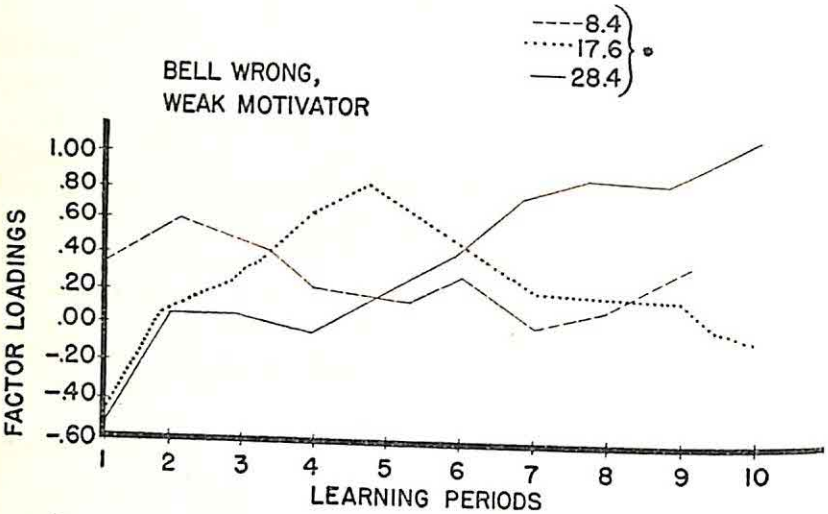


FIG. 5. FACTORS OBTAINED FROM RETHLINGSHAFER'S LIGHT-DARK DISCRIMINATION DATA (43, p. 586).

* Represents the percentage of variance for which each factor is accountable.

(b) producers of less rapid learning: shock-before-choice, food only, jump-before-choice, buzzer-wrong, and buzzer-right. Rethlingshafer employed eight of these motivating conditions, and, as might be expected, obtained the typical three-factor pattern reported by Wherry. Although the patterns are similar, the effectiveness of a given factor does vary under different motivating conditions. This becomes quite apparent from analysis of some of the data reproduced in Figs. 5, 6, and 7. For example, in Fig. 6 where the motivation is food only, no shock, the middle factor is non-operative. On the other hand, this same factor dominates in the rat's learning performance where the motivation is shock-right. Notice in particular the unusually high proportion of the total variance (40.8) which is accounted for by this factor as opposed to zero for performance represented in Fig. 6 and 17.6 in Fig. 5. Such differences are particularly provocative when it is pointed out that the motivator operative in the performance of Fig. 7 was proposed by

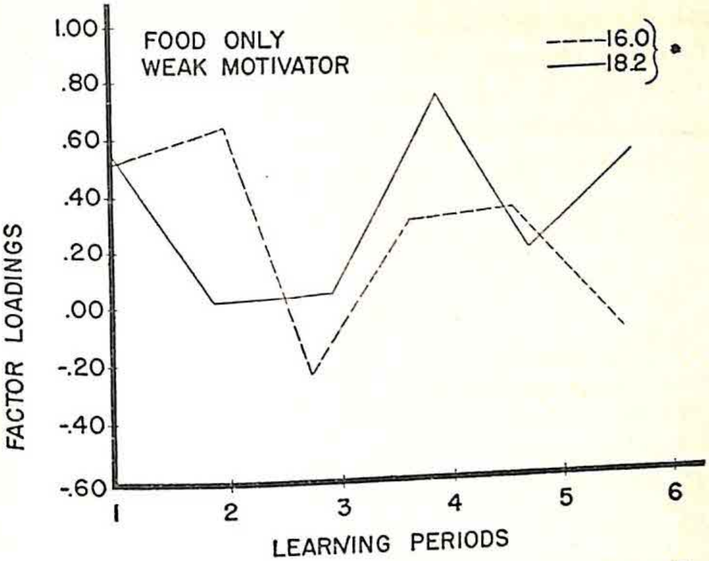


FIG. 6. FACTORS OBTAINED FROM RETHLINGSHAFER'S LIGHT-DARK DISCRIMINATION DATA (43, p. 586).

* Represents the percentage of variance for which each factor is accountable.

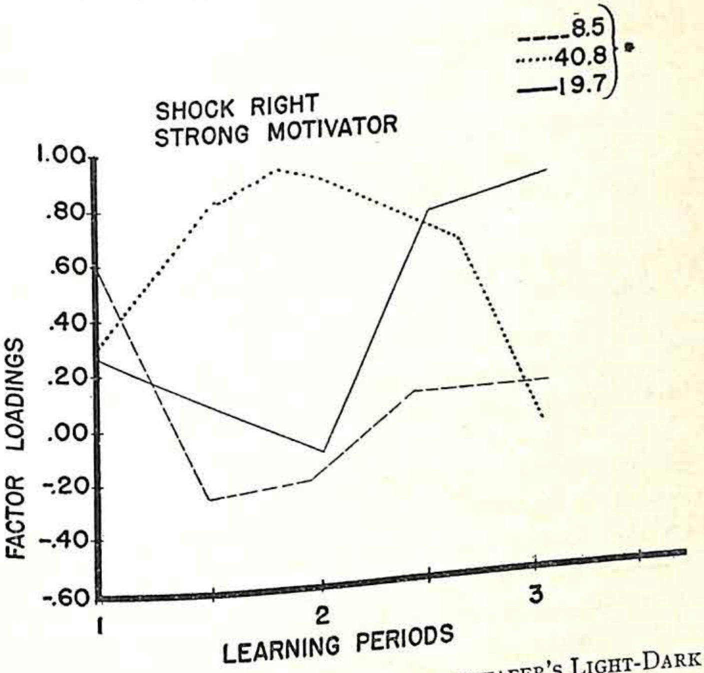


FIG. 7. FACTORS OBTAINED FROM RETHLINGSHAFER'S LIGHT-DARK DISCRIMINATION DATA (43, p. 586).

* Represents the percentage of variance for which each factor is accountable.

Muenzinger as a fast producer of learning, whereas the other two motivators are classified as lesser producers of learning. This type of analysis is, of course, not conclusive. As a matter of fact, in these data there are no indices of consistent trends. The investigator cautions that certain unforeseen aspects of this particular learning problem which resulted in inadequate data actually prevent any legitimate analysis. These limitations are best expressed by the author, who says:

> Under such conditions as shock-right, shock-wrong, delay-at-the-point-of-choice, etc., learning was so rapid that it was impossible to correlate the error scores for all ten days. The number of rats making errors, as well as the total number of errors, had so declined that often by the fifth day the resulting narrow range and skewed distributions made correlations highly unreliable. Hence some matrices were less than 10 by 10, but they represented correlations among day's error scores up to approximately the same state of learning as was represented by a 10 by 10 matrix in a slower learning situation (**43**, pp. 584–5).

Nevertheless, the methodological approach employed is of considerable interest and significance.

EMOTIONALITY-LEARNING STUDY

In addition to the studies reported in the fields of learning, emotionality and motivation, one study has been reported which apparently attempts to cut across all domains of rat behavior. As stated by the authors, Geier, Levin and Tolman (**12**), it is an investigation of the interrelationships between measures of emotionality, hypothesis formation, vicarious trial and error, and visual discrimination learning in the white rat. Using 57 rats, they obtained 29 tetrachoric intercorrelations from two experimental setups, the Hall open field, and a visual discrimination apparatus similar to the Lashley jumping stand. The correlation matrix was analyzed by Tryon's cluster analysis technique. Eleven measures were obtained from the open field test, primarily of an emotional and adaptability nature. The remaining 18 measures were obtained from the discrimination apparatus, primarily vicarious trial and error measures, and hypothesis formation measures. The four factors identified were labelled: open field timidity, cognitive reactivity, unsystematic variability, and non-motivation. The writer is inclined to be particularly critical of this study for two reasons. In the first place, the areas under investigation are "under-determined." This is particularly obvious when one realizes that the entire correlation matrix was derived from two experimental setups. Secondly, the methods of cluster analysis are not as consistent, nor do they result in as psychologically meaningful results as the more stable methods of multiple factor analysis. In essence, of all the studies presented in this survey, the last investigation reviewed is theoretically the most inclusive and methodologically the least adequate.

DISCUSSION

The criticism of underdetermination leveled against the last study can perhaps be directed against virually all of the early investigations presented thus far. Basically, it is the same criticism offered by Wolfle who points out that "the situation is comparable to giving human subjects a dozen or so tests of speed of movement, temperament and emotionality, and learning ability, and then analyzing the resulting correlations. Many low correlations would be expected—and found" (**69**, p. 35). Perhaps what is needed most at the present itme is greater stress on the underlying relationships within a given behavior domain. Then, when the several domains are adequately explored, a giant study covering each of the domains with properly loaded variables, might well be undertaken. Tryon goes so far as to state that

> the scientific study of individual differences is characterized more by the prospects of the future than by a collection of conclusive data . . . in the rat, the major problems of individual differences have been attacked in only a few domains, and many other gross domains of behavior (sense acuities, insight formation, delayed responses, conditioned responses, etc.) remain relatively unexplored. . . . We grievously need more complete studies on more behavior domains in many more species (**62**, p. 360).

If several meaningful parameters can be identified from such batteries, more important studies of the genetic constitution and neurologic functioning of animal "Primary Mental Abilities" could follow. Several well known investigators have expressed considerable regard for such approaches. Wolfle states that "the factorial analysis of ability in animals is still practically untouched as a problem for research. Since animals have many advantages over human subjects for use in future studies of the origin, development, and inheritance of factors of ability, a more thorough knowledge of the factors involved in their ability promises to be of considerable research importance" (**69**, p. 36). More specifically, Price states that "however much the factorist has oversimplified his case in relation to genetical structure, it should be granted that there probably exist genetical sources of intercorrelation in some degree similar to so-called general and group factors" (**41**, p. 184).

On the neurological side, Lashley makes some very pertinent remarks in a paper on the "Coalescence of Neurology and Psychology." So much of what Lashley has to say in this paper is pertinent here that the writer is tempted to include several pages of quoted material. However, the essential relationships between the isolation of functional factors and neurologic functioning is pithily brought out in the following two paragraphs:

> This method is discovering functional variables which are not those of classical psychology. For example, it distinguishes between the ability to think in terms of the spatial relations of objects and to comprehend non-spatial relations. It distinguishes facility in manipulating separate symbols, as in recognition of words in jumbled letters, from facility with combinations of symbols as in grammatical speech. It is difficult to describe these variables except in terms of the tests from which they are derived, for they do not correspond to any familiar classification of functions. They do seem, however, *to correspond to functions which may be independently lost as a result of localized brain injury.*[5] Certain types of apraxia are marked by difficulty in dealing with spatial relations; the function represented by manipulation of isolated symbols resembles the ability which suffers in verbal aphasia as defined by Henry Head, and there are other less clear correspondences.
>
> Psychology has still to discover how the various factors revealed by such analysis interplay to produce organized thought. Neurology likewise still has much to do in the investigation of the interaction of cortical fields which are associated with diverse functions. Nevertheless the discovery that the various capacities which independently contribute to intellectual performance do correspond to the spatial distribution of cerebral mechanisms represents a step toward the recognition of similar organization in neurological and mental events (**34**, pp. 468–469).

The eventual localization of factors within the central nervous system thus becomes a plausible task. Obviously, by controlled destruction of varying areas and varying amounts of the brain and the consequent changes in the factor loadings, perhaps to the extent of complete extinction, important observations can be made relevant to the neurologic nature of such factors. The problem is relatively clear cut, but the first step remains to be accomplished in animal behavior—namely, the identification of stable, meaningful psychological factors.[6]

In the past, critics of factor theory have pointed out that factors are, at most, convenient classificatory conceptualizations. In fact, it has been stated that the nature of a given factor is such that it changes when moved from one battery of variables to another and that this is rather conclusive indication of the mathematically artifactual nature of factors. It has also been pointed out that the number of possible factorial solutions to a given correlation matrix is infinite. In order to meet these criticisms, Professor Thurstone introduced the concept of "simple structure," which assumes that "mind is not a structureless mass, but that it is structured into constellations or groupings of processes that can be identified as distinct functions in the test performances" (**51**, pp. 193–194). Although an hypothesis, this is perhaps the most important conception in the development of factorial analysis as a general scientific methodology. This principle provides a guide in the rotation of arbitrary axes to what Thurstone has called a "psychologically meaningful" solution. Perhaps the most striking thing about this procedure is that it leads to an invariance in the isolation of factors. This phenomenon is not to be taken lightly, since it does not seem reasonable that an event of this sort could occur by chance. The stable factors identified by the rotating principle of simple structure have been referred to as "functional unities," the implication being that these may reflect separate processes which are mediated by the genes, the nervous system, the endocrines, the visual mechanisms, the culture, etc., depending upon the domain under investigation. In particular, it seems reasonable to suspect that "functional unities" of mental ability are a function of the nervous system and its metabolism, and that "functional unities" of emotionality are a function of the physiology of the body as well as conditioning, with genetic mechanisms imposing limits of variability in both domains. The primary contention of this article is that these notions need experimental verification, and that factors involving organismic processes can best be experimentally investigated using animal populations. It is difficult for the writer to see how factor theory can go beyond the behavioral isolation of theoretical parameters (intervening variables, if you wish) as long as it confines itself to human populations. The fact that the results reviewed in this paper have not yet given answers to questions of this type is not an indication of the inappropriateness of factor methodology, but rather reflects insufficient and inadequate utilization of the methods. The point is that it will take a relentless and long-range attack in order to give this approach an apportunity to stand or fall under experimental inquiry. One cannot demonstrate factorial invariance with one factorial study of a given domain any more than one can identify meaningful factors in a battery of ill-conceived variables.

---

[5] Italics mine.

[6] Studies of this type which are aimed at the analysis of hereditary and environmental mechanisms affecting dog behavior are now under way in the Division of Behavior Studies, Roscoe B. Jackson Memorial Laboratory (see **44, 45, 46**).

The writer, then, would defend the factorial approach as the most powerful methodology available on the behavioral level for the unravelling of the ever present mystery of what a given test measures. Comparative psychology cannot afford to ignore the issue of test validity, nor can factor theory afford to stop with the mere identification of factors. A wedding of the two approaches should strengthen each by leading beyond the isolation of unifying concepts to the detailed description of underlying mechanisms of behavior.

SUMMARY

This paper is concerned with the theoretical and experimental implications of animal behavior studies carried out within the framework of factor theory. The writer has reviewed all the known factorial studies of animal behavior, as well as relevant large scale correlation studies. Since only a beginning has been made in each of the areas of learning, emotionality, and motivation, no reliable conclusions can be drawn as to the nature of the psychological factors involved in any of these domains. However, all of the analyses have been consistent in denying the existence of a unitary factor and in pointing up the existence of multiple factors.

The writer has stressed the importance of pursuing these early leads further because of the necessity for verifying experimentally the theoretical concept of "simple structure" on an organismic level. It was also pointed out that the theoretical structure of comparative and phsyiological psychology would be greatly strengthened by a method which could provide insight into the nature of the basic variables involved in its experimental and test situations.

BIBLIOGRAPHY

1. ANDERSON, E. E. The interrelationship of drives of the male albino rat. I. Intercorrelations of measures of drives. *J. comp. Psychol.*, 1937, **24**, 73–118.
2. ANDERSON, E. E. The interrelationship of drives of the male albino rat. II. Intercorrelations between 47 measures of drives and of learning. *Comp. Psychol. Monogr.*, 1938, **14**, No. 8.
3. ANDERSON, E. E. The interrelationship of drives of the male albino rat. III. Intercorrelations among measures of emotional, sexual, and exploratory behavior. *J. genet. Psychol.*, 1938, **53**, 335–352.
4. BAGG, H. J. Individual differences and family resemblances in animal behavior. *Arch. Psychol., N. Y.*, 1920, No. 43.
5. BILLINGSLEA, F. Y. Analysis of intercorrelations between emotionality and other behavior salients in the rat. *Psychol. Bull.*, 1940, **37**, 430.
6. BILLINGSLEA, F. Y. Relationship between emotionality and various other salients of behavior in the rat. *J. comp. Psychol.*, 1941, **31**, 69–77.
7. BILLINGSLEA, F. Y. Intercorrelational analysis of certain behavior salients in the rat. *J. comp. Psychol.*, 1942, **34**, 203–211.
8. CAMPBELL, A. A. Community of function in performance of rats on alley mazes and Maier reasoning apparatus. *J. comp. Psychol.*, 1932, **14**, 225–235.
9. COMMINS, W. D., MCNEMAR, Q., & STONE, C. P. Intercorrelations of measures of ability in the rat. *J. comp. Psychol.*, 1932, **14**, 225–235.
10. CRUTCHFIELD, R. S. Psychological distance as a function of psychological need. *J. comp. Psychol.*, 1939, **28**, 447–469.
11. DUNLAP, J. W. The organization of learning and other traits in chickens. *Comp. Psychol. Monogr.*, 1933, **9**, No. 4.

12. Geier, F. M., Levin, M., & Tolman, E. C. Individual differences in emotionality, hypothesis formation, vicerious trial and error, and visual discrimination learning in rats. *Comp. Psychol. Monogr.*, 1941, **17**, No. 3.
13. Hall, C. S. Emotionality in the rat. I. Defecation and urination as measures of emotionality in the rat. *J. comp. Psychol.*, 1934, **18**, 358–404.
14. Hall, C. S. Emotionality in the rat. II. The relationship between need and emotionality. *J. comp. Psychol.*, 1936, **22**, 61–68.
15. Hall, C. S. Emotionality in the rat. III. The relationship between emotionality and ambulatory activity. *J. comp. Psychol.*, 1936, **22**, 345–352.
16. Hall, C. S. Emotionality in the rat. IV. The relationship between emotionality and stereotyping of behavior. *J. comp. Psychol.*, 1937, **24**, 369–376.
17. Hall, C. S. Drive and emotionality; factors associated with emotionality in the rat. *J. comp. Psychol.*, 1934, **17**, 89–108.
18. Hall, C. S. The inheritance of emotionality in the rat. *Psychol. Bull.*, 1940, **37**, 432.
19. Hall, C. S. Measuring individual differences in aggressiveness in rats. *Psychol. Bull.*, 1941, **38**, 706.
20. Hall, C. S. Temperament: A survey of animal studies. *Psychol. Bull.*, 1941, **38**, 909–943.
21. Hall, C. S., & Martin, R. F. Emotional behavior in the rat. V. The incidence of behavior derangements resulting from air-blast stimulation in emotional and non-emotional strains of rats. *J. comp. Psychol.*, 1941, **31**, 191–203.
22. Hall, C. S., & Klein, L. L. Emotional behavior in the rat. VI. The relationship between persistence and emotionality. *J. comp. Psychol.* 1941, **32**, 503–506.
23. Hall, C. S., & Klein, L. L. Individual differences in aggressiveness in rats. *J. comp. Psychol.*, 1942, **33**, 371–383.
24. Hall, C. S., & Martin, R. F. A standard experimental situation for the study of abnormal behavior in the rat. *J. Psychol.*, 1940, **10**, 207–210.
25. Heron, W. T. The reliability of the inclined plane problem box as a measure of learning ability in the white rat. *Comp. Pscyhol. Monogr.*, 1922, **1**, No. 1.
26. Honzik, C. H. The sensory basis of maze learning in rats. *Comp. Psychol. Monogr.*, 1936, **13**, No. 4.
27. Honzik, C. H. Exteroceptive stimuli and the relative difficulty of maze blinds. *J. comp. Pyschol.*, 1938, **25**, 355–371.
28. Honzik, C. H., & Tolman, E. C. The perception of spatial relations by the rat. *J. comp. Psychol.*, 1936, **22**, 287–318.
29. Hunter, W. S. Studies in the reliability of the problem box and the maze with human and animal subjects. II. Correlation studies with the maze in rats and humans. *Comp. Psychol. Monogr.*, 1922, **1**, No. 1.
30. Hunter, W. S. The behavior of raccoons in a double alternation temporal maze. *J. genet. Psychol.*, 1928, **35**, 374–388.
31. Hunter, W. S. The sensory control of the maze habit in the white rat. *J. genet. Psychol.*, 1930, **37**, 305–357.
32. Hunter, W. S., & Randolph, V. Further studies on the reliability of the maze with rats and humans. *J. comp. Psychol.*, 1924, **4**, 431–442.
33. Husband, R. W. A comparison of human adults and white rats in maze learning. *J. comp. Psychol.*, 1928, **9**, 361–377.
34. Lashley, K. S. Coalescence of neurrology and psychology. *Proc. Amer. phil. Soc.*, 1941, **84**, No. 4, 467–469.
35. Liggett, J. R. A note on the reliability of the chick's performance in

two simple mazes. *J. genet. Psychol.*, 1925, **32**, 470–480.

36. Maier, N. R. F. Reasoning in rats and human beings. *Psychol. Rev.*, 1937, **44**, 365–378.

37. McCulloch, T. L. A study of the cognitive abilities of the white rat with special reference to Spearman's theory of two factors. *Contr. psychol. Theor.*, 1935, **1**, 66.

38. Miles, W. R. The comparative learning of rats on elevated and alley mazes of the same pattern. *J. comp. Psychol.*, 1930, **10**, 237–261.

39. Muenzinger, K. F., Bernstone, A. H., & Richards, L. Motivation in learning. VIII. Equivalent amounts of electric shock for right and wrong responses in a visual discrimination habit. *J. comp. Psychol.*, 1938, **6**, 177–186.

40. Parker, M. M. The interrelationship of six different situations in measurements of emotionality in adult albino rat. *Psychol. Bull.*, 1939, **36**, 564–565.

41. Price, B. Homogamy and the intercorrelation of capacity traits. *Ann. Eugen.*, 1936, **7**, 22–27.

42. Price, B. A note on Professor Tryon's theory of intercorrelation. *Psychol. Rev.*, 1937, **44**, 183–187.

43. Rethlingshafer, Dorothy. The learning of a visual discrimination problem under varying motivating conditions. *J. comp. Psychol.*, 1941, **32**, 583–591.

44. Royce, J. R. A synthesis of experimental designs in program research *J. gen. Psychol.* (in press).

45. Royce, J. R. The mental ability of the dog. *Science*, 1949, **110**, 666.

46. Royce, J. R. A factorial study of emotionality in the dog. (Unpublished Ph.D. thesis, Univ. of Chicago, 1950.)

47. Ruch, F. L. Experimental studies of the factors influencing the difficulties of blind alleys in linear mazes. II. Generalized turning habits. *J. comp. Psychol.*, 1935, **20**, 35–52.

48. Stone, C. P., & Nyswander, D. The reliability of rat learning scores from the multiple T maze as determined by four different methods. *J. genet. Psychol.*, 1927, **34**, 497–524.

49. Thorndike, R. L. Organization of behavior in albino rat. *Genet. Psychol. Monogr.*, 1935, **17**, No. 1.

50. Thurstone, L. L. Primary mental abilities. *Psychometr. Monogr.*, 1938, No. 1.

51. Thurstone, L. L. Current issues in factor analysis. *Psychol. Bull.*, 1940, **37**, 189–236.

52. Thurstone, L. L. Multiple-factor analysis. Chicago: Univ. of Chicago Press, 1946.

53. Tomilin, M. I., & Stone, C. P. Intercorrelation of measures of learning ability in the albino rat. *J. comp. Psychol.*, 1934, **17**, 73–88.

54. Tolman, E. C., & Honzik, C. E. The perception of spatial relations by the rat: A type of response not easily explained by conditioning. *J. comp. Psychol.*, 1922, **36**, 287–318.

55. Tolman, E. C., & Davis, F. C. A note on the correlation between two mazes. *J. comp. Psychol.*, 1924, **4**, 125–136.

56. Tolman, E. C., & Nyswander, D. The reliability and validity of maze measures of rats. *J. comp. Psychol.*, 1927, **7**, 425–460.

57. Tryon, R. C. The genetics of learning ability in rats. *Univ. Calif. Publ. Psychol.*, 1929, IV, 95–115.

58. Tryon, R. C. Studies in individual differences in maze ability, III. The community of function between two maze abilities. *J. comp. Psychol.*, 1931, **12**, 95–116.

59. Tryon, R. C. A theory of psychological components—an alternative to mathematical factors. *Psychol. Rev.*, 1935, **42**, 425–454.

60. Tryon, R. C. Studies in individual differences in maze ability, and a general theory of psychological

components. *J. comp. Psychol.*, 1940, **30**, 283–335.

61. TRYON, R. C. Studies in individual differences in maze ability. VIII. Predictive validity of the psychological components of maze ability. *J. comp. Psychol.*, 1940, **30**, 535–582.
62. TRYON, R. C. Individual differences. In F. A. Moss, (Ed.) *Comparative psychology*, Rev. Ed., New York: Prentice-Hall, 1942. Pp. 330–365.
63. VAN STEENBERG, N. J. Factors in the learning behavior of the albino rat. *Psychometrika*, 1939, **4**, 179–200.
64. VAUGHN, C. L. Factors in the rat. *Comp. Psychol. Monogr.*, 1937, **14**, No. 3.
65. WHERRY, R. J. Factorial analysis of learning dynamics in animals. *J. comp. Psychol.*, 1939, **28**, 263–272.
66. WHERRY, R. J. A test by factorial analysis of Honzik's exteroceptive data. *J. comp. Psychol.*, 1940, **29**, 75–95.
67. WHERRY, R. J. Determination of the specific components of maze ability for Tryon's bright and dull rats by means of factorial analysis. *J. comp. Psychol.*, 1941, **32**, 237–252.
68. WILLIAMS, K. A. The reward value of a conditioned stimulus. *Univ. Calif. Publ. Psychol.*, 1929, **4**, No. 3, 31–55.
69. WOLFLE, D. Factor analysis to 1940. *Psychometr. Monogr.*, 1940, No. 3.
70. WOODROW, H. The relation between abilities and improvement with practice. *J. educ. Psychol.*, 1938, **29**, 215–230.
71. WOODROW, H. Factors in improvement with practice. *J. Psychol.*, 1939, **7**, 55–70.
72. YOSHIOKA, J. G. A further study in discrimination of maze patterns by the rat. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 135–153.
73. YOSHIOKA, J. G. Weber's law in the discrimination of maze distance by the white rat. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 135–184.

Received August 19, 1949.

A NOTE ON THE USE OF THE METHOD OF SUCCESSIVE DIFFERENCES IN EMPIRICAL CURVE FITTING

W. R. GARNER
The Johns Hopkins University

Guilford (1) and later Lewis (2) present a technique called the "Method of Successive Differences" for determining the constants in an equation representing a straight-line relation for a given set of data. The technique essentially involves taking successive pairs of points to get several estimates of the slope constant, and using the mean of these several estimates for the slope constant. Once the slope constant has been determined, the intercept constant is determined by substituting the mean of the X and Y values into the equation and solving for the intercept constant. The purpose of this note is to point out that this method is neither legitimate nor meaningful, and should be discarded from text-book consideration.

Mathematically, the problem is to determine the values of a and b for a given set of data in the following equation:

$$Y = a + bX. \tag{1}$$

With the method of successive differences, the value of b is determined first, as follows:

$$b = \frac{\frac{Y_1 - Y_2}{X_1 - X_2} + \frac{Y_2 - Y_3}{X_2 - X_3} + \frac{Y_3 - Y_4}{X_3 - X_4} \cdots + \frac{Y_{n-1} - Y_n}{X_{n-1} - X_n}}{n - 1}. \tag{2}$$

In many cases the values for X are equally spaced, which means

$$X_1 - X_2 = X_2 - X_3 = X_{n-1} - X_n.$$

Substituting in [2]

$$b = \frac{Y_1 - Y_2 + Y_2 - Y_3 + Y_3 - Y_4 \cdots + Y_{n-1} - Y_n}{(n - 1)(X_1 - X_2)}$$
$$= \frac{Y_1 - Y_n}{X_1 - X_n}. \tag{3}$$

Thus when the values of X are equally spaced, the slope constant is completely and solely determined by two points—the first and the last. The example given by Lewis illustrates this point perfectly. Regardless of what happens to the points in between, as long as the first and last points remain the same, the slope of the "best-fitting" curve will be the same. As pointed out above, the value of a in equation [1] is found by substituting the mean values of X and Y into the equation with the already determined value of b. With this method, it is easy to arrive at some peculiar fits, to say the least.

If the values of X are not equally spaced, then the value of the slope constant is not determined completely by the end values, but to a more or less degree depending on how equally the X-values are spaced. However, in the determination of a slope constant, the weighting of each successive slope estimate should be proportional to the distance along the X-axis. In other words, the calculation should take into account at least the first moment in the weights, or some rather peculiar results can occur. Fig. 1, for example, illustrates how the slope of a straight line can be completely wrong if no account is made of relative distance along the X-axis.

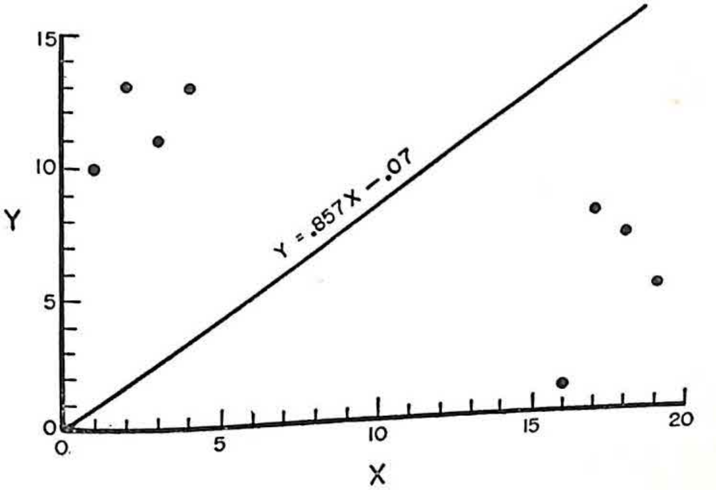


FIG. 1. AN ILLUSTRATION OF A "BEST-FITTING" CURVE DETERMINED BY THE METHOD OF SUCCESSIVE DIFFERENCES.

The straight line shown (and its equation) is the one which best fits the eight plotted points, when the best fit is determined with the method of successive differences.

We have the difficulty then that an erroneous slope can be computed if each slope estimate is not weighted in proportion to the distance along the X-axis. On the other hand, if equation [2] is changed to weight each value correctly, then equation [2] again reduces to equation [3], and once again the slope is determined solely by the two end points.

It is to be hoped that this method will not be perpetuated in future texts.

BIBLIOGRAPHY

1. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.

2. Lewis, D. *Quantitative methods in psychology.* (Lithographed ed.) Iowa City, Iowa, 1948.

Received January 18, 1950

FOUR STUDIES IN PSYCHOLOGY AND SOCIAL STATUS

A SPECIAL REVIEW[1]

RICHARD CENTERS
University of California at Los Angeles

American social scientists in the last two decades have begun to concern themselves increasingly with the social and psychological concomitants of the division of labor in society and the multifarious inequalities that either accompany or stem from it. The commentary and theoretical works of an earlier day, which had only the most meagre of underpinning in direct observation and unchallenged fact, have now given way to the research report based on the community field study, rating scale and other devices of social science. We are rapidly accumulating a voluminous, though perhaps somewhat superficial, body of literature on social status and its significant relationships to personalities and social institutions.

Deserving of a major portion of the credit for the somewhat dubious achievements to date is the anthropologist and sociologist, W. Lloyd Warner, who is senior author of one of the four books to be discussed here and whose concepts and methods are more or less directly reflected in the other three volumes. Eight years have passed since Warner and Lunt published *The Social Life of a Modern Community*, in which they presented the status levels of a New England town as the prototype of the American class structure. Their work, though interesting and valuable, has been characterized by many as somewhat naive in conception and unfortunately shoddy in methodological rigor. A basic fault repeatedly cited in commentary on this and succeeding works of Warner and his many collaborators is the failure of the writers to make explicit and clear some procedure which would make it possible for other social scientists to classify the members of a community in the style practiced by the Warner school. It has been claimed that the criteria of status are too variable, unstandardized and unquantified to be usable by an outsider.

---

[1] WARNER, W. L., MEEKER, MARCIA, and EELLS, K., *Social class in America*. Chicago: Science Research Associates, 1949.

HOLLINGSHEAD, A. B., *Elmtown's youth*. New York: John Wiley, 1949.

HAVIGHURST, R. J., TABA, HILDA, *et al.*, *Adolescent character and personality*. New York: John Wiley, 1949.

STENDLER, CELIA BURNS, *Children of Brasstown*. Urbana: Univ. of Illinois Press, 1949.

THE MEASUREMENT OF SOCIAL STATUS

It is perhaps in a belated attempt to shore up this sagging fundament in a somewhat pretentious edifice that Warner and his two collaborators now give us *Social Class in America*. Save for its first chapter, which delineates the Warner conception of social classes, the book is wholly concerned with the development and use of two techniques for the measurement of social position called Evaluated Participation (E.P.) and the Index of Status Characteristics (I.S.C.).

Of the two, the more basic is the Method of Evaluated Participation, for it is by means of it that communities are originally stratified by the social analyst, and the various individuals in the community allocated to one or another of the social-status levels so derived. In employing it, the status analyst uses one or more of six separate but related devices. They are all evaluative and judgmental, i.e., rating techniques, complementary to one another, but use of one or more of them may suffice to place a given individual more or less accurately in the status scheme. Before using any of them it is first necessary for the investigator to derive what is called a "Social-class Configuration" or "over-all status schema" for the whole community from intensive and prolonged interviews with articulate and status-conscious persons in the community. Such interviews with informants obviously require the best of rapport, and Warner, Meeker and Eells indicate that several such informants are necessary, though they do not reveal what might be acceptable as a numerically adequate sample, saying merely that numerous persons of various social levels are desirable. "A good judge ordinarily is someone who has been in the community all of his life, who thinks about the people of his town in status terms, and who has a vocation that relates him to all social levels."

Granting that this is methodologically adequate for the derivation of an over-all status schema for a community—a point about which more comment later—one is now prepared to place the individuals in the sample in their respective status levels by the various rating devices of E.P. These are:

1. *Rating by matched agreements* of several informants on the placement of many people in the several classes. This is ordinarily accomplished by using the same data employed to construct a composite "Social-class Configuration," since it includes and is largely built upon lists of names of individuals volunteered by informants in their descriptions of the status levels of the community. Apparently, when many informants agree in assigning a given individual to a particular status level, the analyst allocates him accordingly. The degree of agreement required for such placement, unfortunately, remains unstated by the writers, and they appear in discussing the technique of matched agreements to be far more concerned with establishing by it the reality of status levels and status consciousness in the thinking of their informants than in instructing us in its use.

2. *Rating by symbolic placement.* Here an individual is rated by the analyst

as being in a particular social-status level because he is identified by informants with various symbols of superiority or inferiority such as region of residence, social traits and membership character.

3. *Rating by status reputation.* An individual or his family is assigned to a given status level by the analyst because informants say he has a reputation for engaging in activities and possessing certain traits which are considered to be superior or inferior.

4. *Rating by comparison.* In this case the subject or his family is rated as being in a particular status category because informants report him to be equal, superior, or inferior to others whose status level has been previously determined.

5. *Rating by simple assignment.* A person is placed in a given status category because one or more "qualified" informants so assign him.

6. *Rating by institutional membership.* In such case an individual is assigned to a particular status by the analyst because informants report him to be a member of certain institutions which are ranked as superior or inferior. The institutions referred to are families, cliques, associations and churches.

The social-status levels or classes and their membership content derived from these several rating methods constitute for Warner, Meeker and Eells the ultimate class structure of a community. Its validity rests, they say, in effect, in the thinking of their informants themselves—"they are the final authorities about the realities of American social class."

The Index of Status Characteristics is a shorter, more convenient method of measuring social status, offered in the realization that the method of Evaluated Participation is complicated, laborious and time-consuming—and, of course, very expensive. The I.S.C. is computed for each individual by rating him on a seven-point scale for each of four status attributes, namely Occupation, Source of Income, House Type and Dwelling Area, and then multiplying his several ratings by appropriate weightings derived by a regression equation to get a weighted total. The I.S.C. score of an individual is found to predict his social-status level (as determined by the method of Evaluated Participation) in 84 per cent of the cases—for "old Americans." For "ethnics" the percentage of correct placement is smaller, being 72 per cent of the cases.

Warner, Meeker and Eells have on the whole done an excellent job of both carrying through and reporting on what might appear to many as a tedious project. Though apparently the major goal of their work is measuring the social status of people *per se*, it will undoubtedly prove useful in studying a wide range of behavior related to status and class, and no student of these phenomena can affort to neglect it. However, a major flaw, threatening to invalidate the entire superstructure, lies in the inadequacy of the sampling of informants whose opinions are used to construct the over-all social-status schema. The writers point out almost piously that the realities of social class exist in the final analysis in the thinking of people themselves, *which means that social-status* *levels exist as public values, i.e., are a public-opinion phenomenon.* Such being the case, it would seem incumbent upon scientists concerned with accuracy and thoroughness to content themselves with not less than a representative cross-section of informants in building up the basic status schema upon which all else in their system depends. This, however, is not done, and their informants, instead, are a highly selected, highly articulate, higly status-conscious few who appear to be disproportionately representative of our higher social, economic and educational strata. The status levels constructed and the membership in them determined by the methods of Evaluated Participation are thus necessarily dictated by this unrepresentative group. The I.S.C. predicts, also, no more than what would be this group's judgment of an individual's social status. Another serious limitation of the study, of course, lies in its being based upon a single small community, so that the I.S.C. scores presented have meaning only in terms of it.

COMING OF AGE IN THE STATUS SYSTEM

In August B. Hollingshead's *Elmtown's Youth*, the measurement of status functions primarily as a means rather than as an end in itself. Hollingshead has in this work stratified his population into status levels and then gone on to show in what ways and to what extent the status level of the parent conditions the life of his offspring, particularly in the high-school years. To those who are familiar with the Lynds' Middletown books, the Warner group's series of community studies, and others of similar mold, the story told will not be very new, for there is perhaps overmuch in this volume that has been said several times before. Nevertheless, the beastiliness of the American pecking order is sufficiently fascinating to make *Elmtowns' Youth* rewarding reading. Hollingshead has so faithfully reproduced the jungle that we all grew up in and live in that one meets himself or a friend or an enemy on almost any page.

The book is offered as an analysis of the way in which the social system of a Middle Western Corn Belt community organizes and controls the behavior of high-school-aged adolescents reared in it. This analysis is achived by describing the relationships existing between the behavior patterns of 735 such subjects and the positions occupied by their parents in the community's class structure. The seven arenas of social interaction examined are those of the school, the job, the church recreation, cliques, dates, and sex life.

In stratifying the families of his sample, Hollingshead adopted a somewhat roundabout method. He first derived a "control" or reference list of 20 families alleged to belong to one or another of the five status levels which were believed to exist in the community. Then he had a carefully selected group of raters indicate the status levels of the 535 families in the sample by equating their statuses with that of families in the control list. Although it is asserted that "care was taken to select persons from every ethnic and religious group and from every stratum, so that raters would be representative of all sections of the population," it is difficult to believe that the 31 individuals selected were actually adequate for this purpose. Seldom, surely, have so many represented by so few. Probably a more accurate picture of the raters is the following: "Some were in key positions in the institutional structure, others performed functions which brought them into contact with many types of people; and a few were 'plain' citizens. All knew the community through years of living in it, and they were willing to discuss many aspects of community affairs." To this reviewer it is impossible to see this group as "representative of all sections of the population."

Hollingshead is at his best in the sections dealing with the life patterns of the youth of the several status levels, and in describing how the desire for participation, status, self-expression and self-realization in general of working-class youth is systematically frustrated by the power and influence of the ruling class in Elmtown. Often the impact of the data is enhanced by letting the kids tell it in their own words, for they convey, far more poignantly than any objective-minded social scientist allows himself to, the bitterness and unhappiness our social order engenders.

Although *Elmtown's Youth* is especially recommended to American parents by its author, it is unlikely that it will be widely read outside of academic circles, and will more probably be found useful as supplementary reading in university courses in social anthropology and social psychology. If so used, however, it might be well for the instructor to point out that *it is not some disembodied class system that pushes people around in America's Elmtowns, but a flesh-and-blood ruling class and the power it wields.*

CHARACTER REPUTATION AND THE STATUS ORDER

Of the studies under review, probably the most mature, and certainly the most elaborate, is *Adolescent Character and Personality* by Robert J. Havighurst and Hilda Taba, in collaboration with 15 other scientists of The Committee on Human Development of The University of Chicago. The book is a companion piece to *Elmtown's Youth*, being based upon research in the same community and in part on the same boys and girls. But whereas Hollingshead's study was concerned with the behavior of several age groups of adolescents in various spheres of social interaction, this work is an intensive analysis of character in sixteen-year-olds alone. Though concern with the status order is prominently featured in the report, the study is not entirely centered upon status as a character determinant or correlate. Instead, "social-class position" is considered as only one of eight major factors in character formation, the others being school achievement, intelligence, values and ideals, self-adjustment, social adjustment, moral beliefs, and religious activity.

Recognizing that character is a word with many meanings, the authors indicate what they mean by the term more precisely by indentifying it with the phrase "moral character," and further, as "that part of personality which is most subject to social approval." It is added, not too clarifyingly, that having a sense of humor (which is certainly socially approved) is not a part of character but part of that residual something we label "personality."

After conceiving of character as a composite of moral traits, namely, *honesty*, *responsibility*, *loyalty*, *moral courage* and *friendliness*, the next problem, of course, was that of devising some measure of it. Largely on the basis of the monumental studies of Hartshorne and May two decades ago, and the accumulated experience since, behavior was rejected as a usable criterion, and the individual's character reputation was adopted. Reputational ratings for each of the 5 traits as the measure of character were then secured from the individual's teachers, age mates, Sunday-school teacher, school principal, scout leader, and employer. Such judgments were obtained by means of four specially designed instruments, each differing in form, detail and content, but similar in that they described the same general type of behavior manifestation for each trait. Two of them were used with the adult group in the community (teachers, scout leader, etc.) and two with the age mates of the subjects, the object being that of checking on the stability or consistency of the judges' opinions. The mean of the individual's scores on each of the four measures was regarded as the ultimate criterion of his character reputation. Considering the difficulties of measuring such a thing as character, the "reliability" coefficients obtained for the four scales might be considered at least fair at worst, and at best quite good. The lowest was .46 while the highest was a substantial .86. The validity of the instruments is not established.

Adolescent Character and Personality is a rich and meaty offering which contains many challenging leads and insights. But even its authors practically acknowledge that character reputation as they have measured it is, at most, merely an index to the degree to which adolescents conform to the standards of middle-class conduct. They find character reputation to be correlated so slightly with religious activity, moral beliefs, values and ideals, self-adjustment, social adjustment, and even with intelligence and social-class position, that the degree of relationship is of little predictive value. The highest of these correlation coefficients is .52 between social-class position and character reputation, but this is distinctly lower than the correlation of character reputation and school achievement which is .74. Since school achievement, like character reputation, is in a sense a measure of conformity to middle-class standards, it is not surprising that there should be a substantial degree of relationship between it and another measure of conformity.

Although this book makes a contribution to our knowledge of character, its worth as a piece of scientific writing is seriously impaired by an apparently unquestioning adoption and tacit, if not entirely explicit, approval of middle-class patterns of character ideals as the goal toward which character education should strive if the good life in America is to be achieved. It is disturbing to find social scientists so class-centered in their thinking, especially when they themselves recognize that the majority of people are not middle class. Most, moreover, might not want to be.

Children's Awareness of Status and Class Distinctions

Celia Burns Stendler's *Children of Brasstown* is by far the briefest of the research reports considered here, being a monograph of slightly more than 100 pages. But it is by no means the least worthy or interesting, since it is a pioneer thrust into largely unexplored territory. It is a study of "awareness of the symbols of social class" on the part of the children of Brasstown, a New England community of about 15,000 people.

After a brief and quite inadequate "examination" of three other conceptions of social classes, Miss Stendler adopts the Warner definition of class "as a group of people ranked by members of the community in socially superior and inferior positions according to certain socio-economic factors." The social-class position of each of the 107 school children from the first, fourth, sixth and eighth grades who served as subjects was determined by the rating assigned to his parents by a group of five judges. These consisted of three professional men, one businessman and one housewife, who were selected because of their familiarity with the families and for other reasons.

Since Stendler's study is primarily concerned with children's awareness of the *symbols* of social class in the sense of that term previously indicated, one might have expected the author to define the phrase, "symbol of social class." It is disconcerting to find no such definition. Instead, the reader is left to shift pretty much for himself. From the few examples given and the items identified as symbols in the test instruments employed, Miss Stendler seems to mean by a symbol of social class any social attribute or behavioral manifestation that people might use as a criterion of class distinction. Her study involves only a limited number of such characteristics.

The author's procedures in ascertaining the class awareness of her subjects involve a combination of interview and questionnaire techniques in which the approach to consciousness of status and status symbols is often so devious as to yield results of quite dubious significance. The interpretations and conclusions are highly subjective in the main, and at times seem almost gratuitous. First, an interview was held in which children were questioned about their out-of-school activities and their choices of friends as these might indicate class biases, etc. In a second interview they were asked to rate, in economic terms, four sets of pictures illustrating jobs, houses, clothing and recreation which adults had previously adjudged typical of one or another of the several classes. They were also asked to rate themselves and some of their classmates in economic and social terms. Class terms, such as upper-middle or working class, were not employed in such ratings, but instead the children were asked whose parents had more or less money than average, or whose families were high or low in society. This discrepancy in class terminology is unexplained, as is indeed the basis of much of Stendler's other procedure, and it seriously confuses the meaning of her results. A fifth set of pictures was employed in an attempt to find out if, along with increasing awareness of class distinctions, children were also developing associations of good and bad conduct in relation to them. The pictures portrayed children in various acts such as stealing, shooting craps, smoking, fighting, cooperating, studying, sharing toys and returning money found, and subjects were asked whether the child portrayed had little or much money or was high or low in society. A final technique was a *Guess Who* test requiring the child to guess who in his classroom best answered each of 40 brief descriptive questions such as: Whose parent doesn't have a very good job? Who never has any spending money? Who has the nicest home? The object here was to see what relationship, if any, existed between the item in the test and the social-status level of the child mentioned for that item.

In summing up the fruits of her observations (and insights), Stendler characterizes the development of class and status consciousness as progressing through four stages.

> Stage I is that of *Pre-awareness*, and is the condition found to exist for most children below the fourth-grade level, when such terms as 'rich' and 'poor' have little real significance for them.
>
> Stage II, the *Beginnings of Awareness*, extends from the fourth grade to beyond the sixth, and is characterized by signs of increasing class consciousness. The child is apparently more conscious of symbols of lower classness than others, can pick out children with "not much money" in his schoolroom, and can judge them on the basis of their clothes and what they bring to school. He is largely unaware of housing and occupation as class distinctions.
>
> Stage III is that of *Acceptance of Adult Stereotypes*. It begins before grade 6 and continues through grade 8. As the name implies, this is a state of cognitive organization in which the child closely approximates the adult patterns of beliefs and attitudes concerning social classes, including the many contradictions about class which are prevalent in our culture.
>
> Stage IV, *Recognition of Individual Differences*, was achieved by only a few children in the study, and represents a condition of mind wherein the child recognizes the meaning social-status symbols have for most people, but judges people not merely in terms of their class-membership character but in terms of their individual merits as well.

Miss Stendler's delineation of these stages, in the face of her indirectness of method, is quite bold, and of necessity must be accepted with considerable caution until verified in further research. She has, nevertheless, made an important beginning in an area which has hitherto had no systematic exploration.

SUMMARY COMMENT

Though there is obviously in the content of these studies a common bond that ties each of the four into relationship with the concepts of social status and class, each is so discrete in its aim and contribution that an over-all evaluation is rendered difficult for lack of a common problem as a reference point. Each has been of necessity discussed and evaluated in terms of its specific content and achievement, since no central theory or systematic body of principles exists which might define a "Science of Social Classes". The work in this field to date has achieved little in the way of systematization, but consists in the main of exploratory efforts whose results scarcely transcend the descriptive level of discourse. The facts garnered are numerous, but they are not easy to unify into a single theoretical structure. We have anthropological facts, sociological facts, psychological facts, or at least facts gathered by workers in each of these disciplines. And the varied disciplinary affiliation of the facts constitutes in itself something of a barrier to the achievement of an integrated conceptual formulation. We do not yet know whether the systematization of the data we are accumulating will lead to a psychology or to an anthropology or to a sociology of social status and class. The four studies we have been considering make contributions to all three, but their contribution mainly enlarges our store of descriptive material without appreciably contributing towards theoretical advancement in any one of these fields.

A particular conceptual difficulty in most of this research is the lack of adequate differentiation among such concepts as stratification, status and class. These terms are generally used interchangeably and with a breadth of reference that renders them confusingly vague. There has been, however, very noteworthy methodological advance, and with further refinement this may possibly lead to the eventual satisfactory resolution of the conceptual predicament. In Warner's *Evaluated Participation* and his *Index of Status Characteristics*, we are at least moving toward more precise operational definitions of these terms. The Index of Status Characteristics, moreover, holds considerable promise for the facilitation of further research of a somewhat more quantitative and systematic character, since the levels of social status derived from such an index constitute a reasonably objective and meaningful datum with which to work.

Received September 15, 1949.

BOOK REVIEWS

Postman, L., & Egan, J. P. *Experimental psychology.* New York: Harper, 1949. Pp. vii+520. $4.50.

The announced goal of this text is "to give a survey of the main empirical findings and functional relationships in selected areas of experimental psychology, with special emphasis on the control, manipulation, and measurement of variables" (p. xiii). Recognizing the impossibility of covering the field adequately in 500 pages, the authors have limited themselves to the areas of the receptor and perceptual processes (9 of the 20 chapters and 40% of the total pages), learning and retention (6 chapters and 35% of the pages), behavior in social situations (38 pages), reaction time and association (36 pages), analysis of judgment (22 pages), and emotional behavior (18 pages). At the end of every chapter are one or two experiments (30 in all) suitable for the typical laboratory course, designed to illustrate some of the methodological problems described in the chapter. The page or two of accompanying references are arranged in order of the sub-headings of the chapter.

The question naturally arises, Is this an acceptable guide for *the* course called "experimental"? The answer depends first upon one's conception of the pattern of objectives in the undergraduate program. It may be argued that the paramount aim, with either the student of the "liberal arts" or the pre-doctoral candidate, is to impart a realistic and working comprehension of the scientific method, particularly as it is applicable to behavior. Operationally, what does (or, at least, should) a psychologist do when he investigates? What are the premises, the potentialities, the pitfalls, the inherent limitations of his procedure? These insights are not acquired easily, even at the verbal level; it is doubtful if much automatic transfer can be expected from the first chapter of an elementary text or from a detailed description of retinal processes. Perhaps this objective will best be attained by guiding the student through a few representative problems, constantly referring on the one hand to the ideal paradigm of experimental procedure and on the other to some outstanding examples of its application. To this end the second-year course in psychology, particularly, may be dedicated. The survey of general principles can be assumed to be the preserve of the introductory course, and the elaboration of special techniques and the minutiae of findings can be relegated to the more advanced courses. But the backbone of the whole program is this course in Experimental Psychology—a thorough indoctrination into the basics of experiment and a first-hand introduction to the experimental literature of psychology.

If this personal prejudice is accepted, the conclusion will probably emerge that Postman and Egan have not provided the ultimate textbook for such a course. While much research is digested and passed on in the form of somewhat idealized generalizations, too much responsibility is left to the instructor for that most vital task, the critical analysis of the rationale and design of experiments. At the same time, the desire to relieve the student of the "burden" of encyclopedic documentation has been carried too far: there are virtually no references to original sources within the body of the chapters, and the grouped citations are not keyed to the discussion. The result is a highly effective barrier to the student who might otherwise have been tempted occasionally to pursue a point beyond secondary accounts. So our attempt to reach the twin goals of the course receives less than maximal help from the book.

Of more specific evaluation, there will inevitably be much on both sides. Some will disagree with the selection and relative weighting of the major topics. Others will say that considering the great preoccupation with the receptor processes, too little space is devoted to their electrical investigation; receptor theories other than the place and the Young-Helmholtz are worthy of discussion; Helmholtz merits at least one non-hyphenated mention; animal studies of taste preference should be given some space. There will be criticism of a few definitions that somehow fail to come off, and of some statements that might be interpreted to be either descriptions of actual experiments, or logical analyses of problems, or hypothetical formulations, or firmly grounded explanations. Regrets will be expressed over the slightly inconvenient form and location of some tables and figures, and such clerical errors as citing a chapter of Dashiell's in the *Handbook of Experimental Psychology*. But few will question the general workmanship of the book, and many will welcome the straightforward style as a relief from both the closely-written compendium for advanced students and the palsy-walsy primer for beginners.

The final judgment of the book must obviously be a joint function of the book itself and of the purpose for which it is being considered. It will probably find wide use as a combined textual discussion and laboratory manual in experimental courses which are somewhat limited with regard to facilities (laboratory or library), prerequisites, or objectives. In the program discussed above, its most logical adoption would be as a supplementary text for the first-year course, particularly if laboratory work is to be provided.

FRANK W. FINGER.

University of Virginia.

DENNIS, W. *Readings in general psychology*. New York: Prentice-Hall, 1949. Pp. xi+525. $3.75.

Our inflated enrollments in the introductory course have forced us to depend almost entirely upon pre-digested psychology as found in texts and digests of research. Now students in large numbers will have ready access to 71 original reports by authors ranging from Aristotle and Leonardo da Vinci to Neal Miller and Floyd Ruch. Many of the selections are abridgments, but not digests. In size they range from one to a dozen or more pages. The selections are grouped into eleven chapters with conventional titles. Within chapters they are sometimes arranged in logical and sometimes in historical sequence. Each selection is preceded by a brief note which places it in its proper topical setting. Quite often there is a little information about the author and the significance of his contribution.

In approaching the evaluation of an anthology of this kind one is led to wonder how broadly it covers the field and whether its selections are representative of psychological writing and research.

The field covered is broad, the chapter headings being as follows: sight, hearing and other senses, perception, emotion, motivation, learning and retention, reasoning and thinking, intelligence, social behavior, personality, and psychological development. One may question whether sensory and perceptual processes should be given more attention than learning (110 as against 47 pages), but the other fields seem equitably represented.

The teacher who expects to find a book of selections from current research will be somewhat disappointed. Eighteen of the 71 selections were written before 1900. Of the remaining selections, more than one half originated before 1930. Only eight were written within the last decade. Perhaps one reason for such a preponderance of classical and pioneer studies is the fact that so much of recent writing is too technical for ready digestion by beginners. Rather generally the selections appear well-chosen and they should appeal to students as well as to instructors who have not already read them.

Bowdoin College. NORMAN L. MUNN.

VALENTINE, WILLARD L., & WICKENS, DELOS D. *Experimental foundations of general psychology.* (3rd Ed.) New York: Rinehart, 1949. Pp. xxi+472. $3.00.

The readability of psychology textbooks continues to improve. Valentine was a clear and effective writer, and his first chapter, "Magic and Science" is an excellent initial step for the beginning psychology student. Wickens completed the revision of this well-established text by adding and deleting experiments from almost every chapter; completely revising the chapters on Perception, Conditioning, Learning; and adding a new last chapter, Personality. He purposefully omitted a chapter on Social Behavior, a decision, however, with which we do not concur.

The number of sectional and paragraph headings has increased, showing more distinctly the method, results, and summary of the reported experiments. This third revision shows the same trend as the second over the first: more complete integration between the different experiments resulting in a more coherent pattern of subject-matter treatment within each chapter. In contrast to the original edition, the book now closely approximates the typical introductory text with, of course, an emphasis on the experimental findings.

A "revision" usually attempts to improve the presentation within the framework of the original publication. The postwar improvements in textbook printing are seen in the better grade of paper, sharper printing, and more distinct figures and tables. Our chief impression still remains the clarity and interesting style with which the experimental findings are presented and linked together. This effectiveness in presentation was characteristic of Valentine, but the co-author's own contribution is significant. To maintain the contemporary emphasis, the refinements in style, the new content and added comments were introduced without noticeably changing the established pattern of the book. Instructors familiar with two precursors of this new edition will find a better volume, more effectively accomplishing its purpose.

The contrast in content and emphasis between this volume and the introductory text recently written by Guthrie and Edwards, *Psychology: A First Course in Human Behavior*, demonstrates the futility of passing judgment as to which is the "best buy" for use in teaching psychology. Those instructors who prefer the experimental and factual approach are likely to endorse the book under review. If one's teaching of the first course is biased toward principles and applications, this book could best be used only as a reference. This dual purpose as either a text or as a reference book is at the same time its strength and weakness. Many teachers may prefer other introductory texts that attempt to carry the entire load of introducing the student to psychology; or, as references, we may choose other and more specific source materials for showing the experimentalist at work and the factual subject-matter of scientific psychology, providing our own integration and interpretive comment.

While this reviewer does not prefer an experimental and content-laden first course, we recognize and respect the many instructors who do. These individuals will appreciate and profit from the extension and improvements appearing in this latest revision.

STANFORD C. ERICKSEN.

Vanderbilt University.

MURPHY, GARDNER. *Historical introduction to modern psychology.* (Rev. Ed.) New York: Harcourt, Brace, 1949. Pp. xiv+466.

Although it is the last third of this book, outlining the contemporary scene, that differs most from its 1928 predecessor, the minor changes in the earlier parts supply interesting clues to the influence of two active decades on the author's organization and interpretation of the historical trends leading up to the present.

The organization is now quite simple. There are four parts, The Antecedents of Modern Psychology, The Rise of the Research Spirit, Contemporary Psychological Systems, and Some Representative Research Areas—each composed of several chapters. The chapters in turn are broken up into several sections, a device that may disturb the continuity, but helps the reader grasp the writer's plan.

The first edition began with a few paragraphs on the Renaissance. The present one prefixes a brief but illuminating summary of the psychology of a few of the Greek philosophers and even some speculations about primitive psychology.

The thirteen chapters of Part Two that describe the rise of the research spirit have been only slightly altered. British associationism, its origins, variations, and fate, are still treated rather fully. There is the same sympathetic insight into the strange personality and convictions of Fechner, "reaching in one direction and another . . . . bewildered by the complexity of the spiritual heritage," struggling desperately to perceive unity in this life and world. The impact of evolutionism is strongly emphasized. More space has been allotted to William James as new material has come to light. His spirit refuses to die even though, as Murphy puts it, "the research tools which he distrusted have become the chief keys to a technical world for which he had no taste."

History, including psychological history, is often written as if it were the product of a succession of brilliant and vigorous idea men; Murphy, however, when introducing a topic, sketches the philosophical and scientific background, portraying the political and economic culture in which the ideas will grow or die. A quotation will illustrate his plan and his manner. "While behaviorism is expressly a psychological movement, it is an expression of a much more general movement in science (and in the philosophy of science). When it first appeared, it seemed a lonely island; but like all such islands, physical and intellectual, it proved to have many relatives, both visible and hiding."

In Part Three, behaviorism, Gestalt, field theory, and psychoanalysis come up for discussion, in essence at least. The conditioned response is included in the chapter on behaviorism. Other theories of learning and the role of motivation in learning are treated in a separate chapter. It is easy to imagine some learning theorists objecting to the organization of learning theories around associationism and to calling the chapter Modern Conceptions of Association. The chapter on Freud amplifies his later work on ego psychology. The various new movements in psychoanalysis—and their splinters—have a chapter to themselves with the title The Response to Freud. Each system receives a sympathetic hearing, with only occasional criticism.

Part Four attempts to bring certain areas of research up to date:

intelligence, physiological psychology, child psychology, social psychology, and personality. It ends with a few pages of general interpretation. This is the part which will receive the most criticism. The author anticipates this by noting that the selection of topics is a personal one, that he has omitted comparative psychology, several well developed areas of applied psychology, abnormal psychology, and others. He regrets that he has not digested the military research—even in social psychology—carried out during World War II. Although Murphy has often stressed the systematic importance of motivation and the first edition had a chapter on instinct, this topic now has no chapter, nor section of a chapter, to itself.

No one, to be sure, could balance a book like this one so as to bring all readers' interests into equilibrium. Most psychologists will find almost everything they have a right to expect treated here, and treated with respect as well as depth. Graduate students and college seniors—if they are not dismayed by occasional sixty-word sentences—will find clarity as well as breadth.

DONALD M. JOHNSON.

Michigan State College.

HUMPHREY, G. *Directed thinking.* New York: Dodd, Mead, 1948. Pp. 229. $3.50.

HUTCHINSON, E. D. *How to think creatively.* Nashville: Abingdon-Cokesbury Press, 1949. Pp. 237. $2.75.

Gradually we psychologists learn more of the social responsibility entailed in our roles as scientists. One way of fulfilling this obligation, and at the same time preventing the prostitution of our science by the quack, is in writing the results of our researches so that they may be understood by our fellow citizens. Both Humphrey and Hutchinson have performed this mission in their books about the reasoning and creative thinking process.

From the point of view of his fellow scientist, Humphrey's work is a paragon in popularization. He cites much of our experimental literature as proof for his assertions; he constantly reminds his reader of the fundamental tenets of our scientific approach. For instance, in using the results of experiments by Ach, Maier, Thorndike, and others, he skillfully conveys the relevance of these researches to his reader by interweaving them with concrete, practical examples from everyday experience. In discussing the erroneous nature of the trial-and-error theory, Humphrey manages to help the reader understand that such a theoretical mistake in itself may be helpful, "because it has brought the facts into the open." Humphrey broadly conceives the process of "directed thought" or reasoning (as contrasted with "free thought" or reverie) as including such aspects of thinking as imagery, insight, concept formation, symbolism, and generalization. The wide coverage at times induces in the reader an uncomfortable sense of hodge-podge. This lack of integration is assuredly more the result of the state of psychology than of Humphrey's skill as narrator. The psychologist, as scientist, however, will look in vain for new interrelations, which might have appeared in the course of Humphrey's evaluation of literature in this area.

Hutchinson's work will do little to help laymen gain appreciation for the use of more modern, experimental, scientific research techniques in such difficult areas as creative thinking. Although he stresses the importance of the emotional and unconscious factors which enter the process of thinking, his own methodology is infected by a late 19th century introspectionism. He bases his account upon the results of a questionnaire on creative effort given to some 250 notable American and British thinkers, (e.g., Aldous Huxley, Arnold Bennett, Bertrand Russell), rich anecdotal reports he has collected from magazines and books, and his own conversations and thinking with engineers and inventors. Hutchinson neglects the valuable supports he might have gained from the scientists whose techniques allow experimentation with those unconscious factors which are barred to introspective techniques. Toward the end of his work, his weak faith in the scientific method is revealed in such an assertion as, "Creative thought . . . is the history of these factors in multiple and ever-changing patterns. For this reason no one will ever write the whole account, for it is written not in words but in unfathomable experience" (p. 196).

Throughout Hutchinson's book one senses he is not writing merely for the layman. Starting this work when he was studying at Cambridge with McDougall and Bartlett in the early twenties, Hutchinson makes a masterful attempt to integrate his considerations about the creative process, both in its logical and intuitive facets. His result is far superior to Wallas' stages of creative thinking. Yet, it does not result in a theory readily subject to experimental testing. Only part of Hutchinson's success in unifying his treatment, as contrasted with Humphrey's, comes from his exclusion of findings in such tangential areas as concept formation; the other part of his success is due to his far greater willingness to make assertions without supporting experimental facts.

There are important differences between the two books in the extent to which they are permeated by the scientific as contrasted with the philosophic approach. Both writers have done a very real service to the lay reader in conveying the impact which the more recent emotional-motivation emphases in psychology have had upon treatment of the "higher" mental processes. Both books stress the non-logical, unconscious nature of the psychological process even in such a stronghold of pre-Freudian psychology as thinking.

University of Michigan. HAROLD GUETZKOW.

SWENSON, ESTHER J., ANDERSON, G. LESTER, & STACEY, CHALMERS. *Learning theory in school situations.* Minneapolis: Univ. of Minnesota Press, 1949. Pp. 103.

This monograph should be of interest to the practical-minded educator as well as to the student of learning theory. It includes abstracts of three doctoral dissertations on the psychology of learning in the schoolroom situation, and an introduction which attempts to extract from these studies certain theoretical implications.

The findings, many of which have practical significance for instructional purposes, are too numerous to be recounted in this brief review. Dr. Swenson, working with second grade pupils, reports a highly significant superiority of a "meaning" or "generalization" method of teaching simple facts of addition as compared to two "drill" methods. This superiority was especially evident when a transfer criterion of learning was used.

Dr. Anderson applied similar type methods in teaching fourth grade students arithmetic skills. When the criterion of learning was response reproduction—i.e., the pupils had to recall specific knowledge they learned—". . . there was little difference between the pupils who had relied on repetition of specific elements and those who had followed procedures which encouraged understanding and meaningful organization. But when the criterion was the ability to apply what presumably had been learned to new situations . . . the pupils who had been encouraged to comprehend mathematically what they were learning and to capitalize meaningful relations were significantly superior." This, however, was not true for all the students; pupils of relatively low ability and good achievement learned better under the drill method while students of relatively high ability and poor achievement learned better under the meaning method.

Dr. Stacey tested sixth grade children in their ability to solve certain verbal problems (select the word that did not fit in a group of five words). Five methods of training were used, each varying in the degree of information given to the subject. The results were that "active participation involving self discovery" was superior to "passive participation involving recognition . . . of information previously provided. . . ." The results of this study of problem solution tended to support "Thorndike's conclusions concerning the efficacy of reward and the inadequacy of punishment"—especially when the self-discovery technique was used.

The theoretical implications of these three experiments are integrated in the introduction written by Dean T. R. McConnell of the University of Minnesota. Quoting extensively from Hilgard's *Theories of Learning* to support his own analysis, McConnell considers the initial two studies of the monograph as being a test of the "relative efficacy of learning" according to the association and field conceptions of learning.

The validity of this analysis rests upon the premise of the *existence* of *a field theory* and *an association theory* capable of generating deductions relevant to the experimental situations of Drs. Swenson and Anderson. This reviewer does not think this to be the case. At best, the terms "field" and "association" have denotative meaning only if they refer to formulations which certain writers for various reasons group together. The connotative meaning of these terms, which vary markedly from individual to individual, should not be confused with scientific theorizing. In this monograph, association theory is equated with some statements from Thorndike's unsystematic writings. The evidence creates serious doubt about the wisdom contained in such statements. This, however, does not justify the generalizations made to association theories as a group. Variations in motivation and reward which play such an important role in some of the association theories are almost completely ignored in the experimental aspects resulting from the "meaning" type of pedagogy. Nor is any cognizance expressed of Hull's statement about the practical educational implications of his associationistic theory: "On the basis of the preceding principles it seems reasonable to suppose that if children from an early age were systematically trained to find the solution of genuine individual problems by means of their own symbolic processes, intellectual education might be far more effective than it is at the present."[1]

The lack of rigor of the theoretical interpretation should not detract from the well-executed, "tough-minded" research contained in this monograph. It merely emphasizes the unfortunate hiatus which exists between practical minded educators and learning theorists.

New York University. HOWARD H. KENDLER.

MEAD, MARGARET, *Male and female: A study of the sexes in a changing world.* New York: Morrow, 1949. Pp. xii+477. $5.00.

In this new volume, Margaret Mead gives us the benefit of a quarter of a century's observations on male and female in a variety of cultures. Her purpose in writing the book is "to give to the reader the positive findings from a comparative study of culture about similarities, about the essentials in maleness and femaleness with which every society must reckon, and regularities as well as differences" (p. 32). Her emphasis throughout is biosocial and provides a corrective to her previous apparent overplaying of cultural at the expense of biological factors.

By way of introduction Dr. Mead takes the reader on a verbal field trip and by a succession of concrete word pictures permits him to share at "first-hand" some of her own experiences in the South Seas. The three remaining sections constitute the main body of the book. Part II, "The ways of the body" is a child's-eye view of the varying ways in which he discovers and assumes his sex role in the seven primitive societies that the author has intensively studied. Part III, "The problems of society," is a reciprocal of Part II, revealing the varying ways in which these societies have patterned the sex roles of their people. Finally Part IV brings the author's ethnological insights to bear on "The two sexes in contemporary America." Following the presentation of the central material there are four supplementary sections which elaborate the points in the text and provide sufficient documentation to compensate for its meagerness in the author's earlier popular works.

---

[1] Hull, C. L. *Conditioning: outline of a systematic theory of learning in the psychology of learning.* 41st Yearbook, Part II, National Society for the Study of Education. Chicago: Univ. of Chicago Press, 1942.

Whatever the reader's bias, he can hardly fail to recognize *Male and Female* as a landmark in a chaotic field, providing a new frame in which to view old problems. The unique feature is the handling of each problem by means of direct intercultural comparisons,—a treatment conducive to objectivity. According to Dr. Mead, an important lesson preliterate peoples can teach us is that underlying wide diversity in superstructure, all cultures share certain basic regularities in biosocial development which differentiate male and female the world over. As a result, the author sees security for the male to depend on achievement; that for the female, on nurturing. The moral is for the good society to take account of these fundamentals in assigning roles to its men and women, while at the same time permitting sufficient freedom of choice to allow full expression of the abilities of both sexes. Having taken this stand we are surprised to find the author in sudden retreat, in an attempt to escape from what she fears may be too much freedom for women, threatening both their men and their femininity. It would be more in character for Dr. Mead to be seeking ways of channeling the freedom she fears into new, more cooperative patterns of behavior between the sexes that would remove the sources of threat to both.

Although the viewpoint expressed in *Male and Female* makes interesting reading, the scientist will be chiefly concerned with the validity of its basic assumptions. Some of these we find rooted in the hypothetical psychodynamics of the preverbal infant. "So the female child's earliest experience is one of closeness to her own nature. The little boy, however, learns that he must begin to differentiate himself from this person closest to him" (p. 148). In view of the aheuristic nature of such "earliest experiences," it would seem extremely risky to use them as the cornerstone of the New Society. There are other questionable assumptions which, though testable, have not yet been put to test. An example concerns women's nurturing functions, "In the case of women, it is only necessary that they be permitted . . . to fulfil their biological role, to attain this sense of irreversible achievement" (p. 160). Finally, there are those assumptions that have neglected the "negative instances" for which the author herself specifically alerts the scientist.

Thus we find that the keynote of the chapter on "Rhythm of work and play" is the notion that women are subject to periodic fluctuations in capacity, extensive research literature to the contrary notwithstanding! Throughout the volume, the reviewer had the uncomfortable feeling that evidence was sacrificed to theory, and often wished Dr. Mead had depended less on "the eye as a scientific instrument" and had actually gotten down to the counting of neckties and other things she felt so free to dispense with. So much reliance on intuition leaves the reader wondering whether the picture portrayed in *Male and Female* may not be a little "artistically out of focus."

Although a brief review cannot touch on the many stimulating ideas advanced in Dr. Mead's present work, her treatment of "The ethics of insight-giving" which she tucks away in an appendix is too important to ignore. Here she defends her conspicuous omission of social pathology on the ground that "Society, the patient" might misuse the insights and become the worse for them, as she fears may have resulted from the dissemination of the Kinsey report. It seems to the reviewer that this position involves not only the question of the ethics of insight-giving but also the question of the ethics of censorship in a democracy. Unless we accept a dictatorship of experts, society will have to be therapist as well as patient and will need all the insights it can get to cure itself. The cultism implied in Dr. Mead's arguments for witholding information seems wholly irreconcilable with her own plea for freedom through knowledge. What seems to be indicated is *more* knowledge so that *more* insights may become available, for freedom from social no less than from individual disease in a democracy must depend on an enlightened public.

From cover to cover, *Male and Female* is provocative, raising more questions than it can answer. Its chief value will be to the social philosopher but the social scientist will also find it a rich source of hypotheses. In fact, as long as the concepts presented are recognized as hypothetical rather than conclusive, this book can make a constructive contribution to our thinking in a controversial field.

The University of Southern California. GEORGENE H. SEWARD.

LINDESMITH, A. R., & STRAUSS, A. L. *Social psychology.* New York: Dryden Press, 1949. Pp. xvi+549. $4.50.

This introductory text is offered as an exposition of the point of view that human behavior is profoundly affected by the facts that people use language and live in social groups. The book deals much more with language than with social groups, although the two are linked by showing that social organization and culture depend on language, and that socialization is mediated by language.

The first quarter of the book is a discussion of language behavior.

This section is devoted largely to pointing out ways in which language differentiates men from other animals, and with defining and describing various aspects of language behavior—symbols, symbolic environments, perception, remembering, etc. Emotional behavior is included as an aspect of language, though physiological aspects are recognized.

The second quarter of the book discusses socialization. The reader may be surprised to find that, of the five chapters offered, one deals with language development in the child and two with thinking as dependent on language and on group memberships. The other two chapters briefly discuss social roles and "the self," again with a strong emphasis on language aspects. The self is defined behavioristically as a set of regulatory responses, which results from internalizing the regulatory social responses of others.

The third quarter of this text is on personality. What unity it possesses depends on the fact that attitudes, traits, conflicts and breakdowns are related to social roles, while mechanisms such as rationalization and projection are treated as "symbolic adaptations" (Ch. 11). The chapter on deviant behavior, dealing with feeblemindedness, neuroses and psychoses, seems hard to justify in this text.

The final quarter of the book gives an elementary and rather sparse discussion of race and sex differences, and of group behavior and social change.

In this reviewer's opinion, this text fails to present an integrated account of social behavior. Most fundamental, it is largely descriptive and lacks any basic theoretical orientation. Much of the discussion is purely topical. Throughout the book, it is asserted that motives, attitudes, language, roles, etc. are learned. Yet learning theory and research are dismissed with the statement that "these materials have not proved especially illuminating or helpful for the social psychologist" (525). Freudian theory is disposed of in a section on "misconceptions of child behavior" (431). Lewinian theory is not considered. It is significant that such terms as sociometry, frame of reference, and group dynamics do not appear. Some use is made of Mead's point of view in discussing roles and "the generalized other" (175–178).

This text makes frequent use of research findings to support the discussion. However, there is no description of research methods in social psychology beyond three pages on personality tests. While the authors promise in the Preface to point out research implications growing out of their point of view, they rarely do this. The text includes sections criticizing the naïve use of such terms as mind, self, instinct, etc., reflecting the biological-behavioristic tradition.

The volume is written with marked simplicity and clarity, with pertinent and interesting quotations and illustrations. Students who have previously had an introduction to general psychology will find much repetitive material in this text. Superior undergraduates are apt to regard this text as too elementary. Instructors who wish to introduce their students to social groups and personality organization, and to the theories and research methods being developed to measure and relate them, will probably find that this text does not meet their purposes.

GLEN HEATHERS.

Fels Research Institute.

FRANK, JEROME. *Courts on trial: Myth and reality in American justice.* Princeton: Princeton Univ. Press, 1949. Pp. xiv+441. $5.00.

This is an analysis of judicial justice and injustice in America. In fact, the subtitle would have been a better title, namely, *Myth and Reality in American Justice.*

Basically the author describes the *real* or *true* workings of our courts with which most people are not acquainted. The result is a significant study of many of the myths and legends about the law and its functioning, as contrasted with the way it really operates. The theme of "modern legal magic" recurs several times.

Judge Frank deals realistically and objectively with one of our most important social institutions. He writes, "(1) of what courts actually do, (2) of what they are supposed to do, (3) of whether they do what they're supposed to do, and (4) of whether they should do what they're supposed to do."

The result is an intriguing explanation of the uncertainty of much of our legal lore, as contrasted with the supposed certainty of the law stressed by a great many lawyers. The author analyzes in considerable detail the importance of personal factors in the functioning of the law—as it operates through human beings called lawyers, witnesses, jurymen, and judges. He shows how "court-house government" often obstructs the very justice it seeks to secure.

Judge Frank has had a distinguished career as lawyer, Chairman of the Securities and Exchange Commission, and more recently as Federal Judge; but he is perhaps better known to scholars for his own scholarly writings. All of his books and many of his law review articles have psychological themes; and in the present instance one is reminded again and again of his *Law and the Modern Mind*, published twenty years ago.

With the possible exception of the late Professor Edward S. Robinson's *Law and the Lawyers*, the best appraisals of the psychology of law have been written by lawyers, such as Leon Green's *Judge and Jury*, Huntington Cairns' *Law and the Social Sciences*, and Joseph N. Ulman's *A Judge Takes the Stand*. The present work demonstrates both in its writing and documentation the author's vast amount of reading and study in many different phases of the law, as well as in psychology sociology, anthropology, and psychoanalysis.

However, in spite of Judge Frank's hope that the book will be useful for "intelligent non-lawyers as well as lawyers," it is doubtful whether it will have any special appeal to those without formal education in the law. Max Radin's *The Law and Mr. Smith* is essentially the kind of book which most psychologists would find more useful in understanding in broad terms what the law is about. Frank's volume, on the other hand, is after all really written for lawyers—and, it might be said, only for that small percentage of lawyers who are essentially legal scholars. At least, only those persons with considerable legal training are likely to be interested in reading the volume in detail.

STEUART HENDERSON BRITT.

McCann-Erickson, Inc., New York City.

JOHNSON, PALMER O. *Statistical methods in research.* New York: Prentice-Hall, 1949. Pp. xvi+377. $5.00.

Several characteristics of this book stand out. The book is primarily an exposition of the methods, techniques, and approaches to statistical problems developed by R. A. Fisher and others. The purpose of the book is not to present a compilation of formulas and derivations. Few formulas are derived, but sources are cited so that the more advanced students may easily find the original papers. The real purpose seems to be that of presenting statistical techniques as research tools; showing how they are applicable, under what conditions each should be used, and permissible ways of interpreting the computed statistics. The emphasis of the book is on the interpretive rather than the descriptive function of statistics. "Thinking statistically is equivalent to thinking scientifically."

The book was written as a text book for a year course for students who have had some training in statistical techniques. While it is claimed the book was written for readers without *specialized* mathematical training, it must be noted that this does not imply "*without* mathematical training." For text book use, each chapter is well equipped with problems and examples to be solved. The examples used to illustrate the principles presented are drawn from real problems, and hence add to the realism of the concepts presented. The problems for the student to solve have more reality than is customary in statistical text books.

It is difficult to isolate any few chapters as more outstanding or better written than others. The chapters primarily concerned with the testing of statistical hypotheses will probably contribute more to the thinking of graduate students than do some of the other chapters. For other readers, some of the other chapters will be more valuable because of their "net" contribution to their thinking and to their techniques.

The book is valuable as a reference book, a source where one may find a variety of statistical functions with more than their mathematical expression—with information regarding the applicability and interpretation of such functions. Five useful tables are included at the end of the book. These are tables of normal distribution, t, χ^2, F, and L_1.

HAROLD A. EDGERTON.

Richardson, Bellows, Henry & Co., Inc.

BOOKS AND MATERIALS RECEIVED

Ackerman, Nathan W., and Jahoda, Marie. *Anti-Semitism and emotional disorder*. New York: Harper, 1950. Pp. xiv+135. $2.50.

Adorno, T. W., *et al. The authoritarian personality*. New York: Harper, 1950. Pp. xxiii+990. $7.50.

Bettelheim, Bruno, and Janowitz, Morris. *Dynamics of prejudice*. New York; Harper, 1950. Pp. xix+227. $3.50.

Boss, M. *Meaning and content of sexual perversions*. New York: Grune & Stratton, 1949. Pp. viii+153. $3.50.

Bryngelson, Bryng, Chapman, Myfanwy E., and Hansen, Orvetta K. (Rev. Ed.) *Know yourself*. Minneapolis: Burgess Publishing Co., 1950. $2.00.

Cohn, Robert. *Clinical electroencephalography*. New York: McGraw-Hill, 1949. Pp. 639. $14.00.

Dennis, Wayne. (Ed.) *Current trends in industrial psychology*. Pittsburgh: Univ. Pittsburgh Press, 1949. Pp. 198. $3.75.

Derber, Milton. (Ed.) *Proceedings of the first annual meeting*. Vol. I, No. 1. Champaign, Ill.: Industrial Relations Research Association, 1948. Pp. ix+225.

Eisenson, Jon. *Basic speech*. New York: Macmillan, 1950. Pp. xi+344. $3.00.

Engle, T. N. *Psychology*. New York: World Book Co., 1950. Pp. xi+628.

Epstein, Bertram. *Immediate and retention effects of interpolated rest periods of learning performance*. New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949. Pp. xiii+77. $2.10.

Feldenkrais, M. *Body and mature behavior*. New York: International Universities Press, 1950. Pp. viii+167. $3.75.

Freeman, Frank. *Theory and practice of psychological testing*. New York: Holt, 1950. Pp. xxi+518. $3.50.

Glover, Edward. *Psychoanalysis*. New York: Staples Press, 1949. Pp. 367. $4.00.

Greber, Leo. *Production of new housing*. New York: Social Science Research Council, 1950. Pp. ix+186.

Guilford, J. P., and Michael, William B. *The prediction of categories from measurement, with applications to personnel selection and clinical prognosis*. Beverly Hills, Calif.: Sheridan Supply Co., 1949. Pp. 55. $1.40.

Haas, Robert Bartlett. *Psychodrama and sociodrama in American education*. New York: Beacon House, 1949. Pp. xiii+25. $6.75.

HARMS, ERNEST. *Awakening into consciousness of sub-conscious collective symbolism as a therapeutic procedure.* New York: Child Care Publications, 1949. Child Care Monograph No. 2. $1.00.

HARSH, CHARLES M., AND SCHRICKEL, H. G. *Personality.* New York: Ronald Press, 1950. Pp. vii+518. $5.00.

HEBB. D. O. *Organization of behavior.* New York: Wiley, 1949. Pp. xix+164. $2.50.

JACKSON, LYDIA, AND TODD, KATHLEEN M. *Child treatment and therapy of play.* (2nd Ed.) New York: Ronald Press, 1950. Pp. xiii+159. $2.50.

JOHNS, RALPH LESLIE. *Psychology in everyday living.* New York: Harper, 1950. Pp. x+564. $3.50.

JOSEY, CHARLES C. *Psychology and successful living.* New York: Scribner's, 1948. Pp. xiv+405.

KALLEN, HORACE, *The education of free men.* New York: Farrar, Straus, 1949. Pp. xix+332. $5.00.

KATZ, DAVID. *Gestalt psychology.* (Tr. by Robert Tyson.) New York: Ronald Press, 1950. Pp. x+175. $3.00.

KORNHAUSER, ARTHUR. (Ed.) *Psychology of labor-management relations.* Champaign, Ill.: Industrial Relations Research Association, 1949. Bulletin No. 3. Pp. vi+122. $1.50.

KUBIE, LAWRENCE S. *Practical and theoretical aspects of psychoanalysis.* New York: International Universities Press, 1950. Pp. xvii+252. $4.00.

LEVINSON, HORACE. *The science of chance.* New York: Rinehart, 1950. Pp. 348. $2.00.

LOWENTHAL, LEO, AND GUTERMAN, NORMAN. *Prophets of deceit.* New York: Harper, 1949. Pp. xiii+164. $2.50.

MASSING, PAUL. *Rehearsal for destruction.* New York: Harper, 1949. Pp. xviii+341. $4.00.

MEERLO, JOOST A. M. *Patterns of panic.* New York: International Universities Press, 1950. Pp. 120. $2.00.

MOLONEY, JAMES CLARK. *The magic cloak.* Wakefield, Mass.: Montrose Press, 1949. Pp. xvii+345. $5.00.

O'BRIEN, PATRICK. *Emotions and morals.* New York: Grune & Stratton, 1950. Pp. xiii+241. $3.50.

OHWAKI, Y. *Tohaku psychologica folia.* Tohaku University: Fendai, Japan, 1949. Pp. 69.

REIWALD, PAUL. *Society and its criminals.* New York: International Universities Press, 1950. Pp. x+315. $4.50.

ROBACK, A. A. *Personality in theory and practice.* Cambridge: Sci-Art, 1949. Pp. 437. $3.75.

ROMANO, JOHN. *Adaptation.* Ithaca: Cornell Univ. Press, 1949. Pp. xiv+113. $2.00.

RUBIN. EDGAR. *Experimenta psychologica.* Copenhagen: Ejnar Munksgaard, 1949. Pp. 355.

SENN, MILTON J. E. (Ed.) *Problems of infancy and early childhood.* New York: Josiah Macy, Jr., Foundation, 1949. Pp. 150.

SEQUIN, C. ALBERTO. *Introduction to psychosomatic medicine.* New York: International Universities Press, 1950. Pp. 320. $5.00.

SONNEMANN, ULRICH. *Handwriting analysis.* New York: Grune & Stratton, 1950. Pp. x+276. $5.00.

SPURLING, R. GLEN. *Practical neurological diagnosis.* (4th Ed.) Springfield, Ill.: Charles C Thomas, 1950. Pp. xx+268. $5.00.

STECKLE, LYNDE C. *Problems of human adjustment.* New York: Harper, 1949. Pp. x+350. $3.00.

STEINER, M. E. *The psychologist in industry.* Illinois: Charles C Thomas, 1949. Pp. vii+107. $2.00.

TOLMAN, E. C. *Purposive behavior in animals and men.* Berkeley: Univ. California Press, 1949. Pp. xvi+463. $5.00. (Reprinted.)

TREDGOLD, R. F. *Human relations in modern industry.* New York: International Universities Press, 1950. Pp. 192. $2.50.

TSANOFF, RADOSLAV A. *The ways of genius.* New York: Harper, 1949. Pp. xv+310. $3.50.

TURNBULL, JOHN A. *Labor-management relations.* New York: Social Science Research Council, 1949. Bulletin No. 61. Pp. ix+112. $1.25.

VALENTINE, P. F. (Editor). *The American college.* New York: Philosophical Library, 1949. Pp. xvi+575. $10.00.

WALLIN, J. E. WALLACE. *Personality maladjustments and mental hygiene.* (2nd Ed.) New York: McGraw-Hill, 1949. Pp. xiv+581. $5.00.

WATKINS, JOHN G. *Hypnotherapy of war neurosis.* New York: Ronald Press, 1949. Pp. x+384. $5.00.

WEISS, EDOARDO. *Principles of psychodynamics.* New York: Grune & Stratton, 1950. Pp. ix+268. Pp. ix+268. $4.00.

WORMHOUDT, ARTHUR. *The demon lover.* New York: Exposition Press, 1949. Pp. 150. $2.50.

WOODWORTH, HUGH. *The nature and technique of understanding.* Vancouver: Wrigley Printing Co., 1949. Pp. 142. $4.00.

Inventory of research in racial and cultural relations. Chicago: Commission on Education, 1949. Vol. II, Bulletin 2. Pp. 67.

Vol. 47, No. 4 July, 1950

Psychological Bulletin

SOME PSYCHOLOGICAL PROBLEMS OF ECONOMICS

SAMUEL P. HAYES, JR.
Department of State[1]

I. INTRODUCTION

It is the practice these days to genuflect before the ideal of interdisciplinary cooperation, as a prerequisite to progress toward the distant goal of a single, basic, integrated social science. Most discussions of the interrelationships of the social sciences emphasize the need for cooperative research, and philanthropic foundations have abundant funds to support projects that bring together the different social sciences. A great impetus to such cooperation, moreover, comes from the practitioners of applied social science; faced with the concrete problems of government, business, education, diplomacy, the social science technician quickly finds that his recommendations are no more reliable than the weakest links in the chains of reasoning supporting them. Usually trained in only one, or at most two, social sciences, the social scientist on the firing line is forced to look for help from others whose training will supplement his own.

Yet, for all the good intentions held and expressed, far too little cooperative social science is practiced. Partly, this is an inevitable result of the great accumulation of social science knowledge, making narrow specialization the obvious route to expertness and advancement. Partly, it comes from a tradition of cloistered and solitary scholarship, and from the personality structure of many who find such scholarship congenial. And partly, it comes from interdepartmental jealousies in universities, business enterprises, and government. But there remain many social scientists with an interest in and a competence for cooperative, interdisciplinary research. The present paper has been prepared to suggest some of the many possibilities for fruitful collaboration between economists and psychologists in the study of economic behavior. It considers primarily the needs of economics for psychological supplementation, recognizing at the same time that cooperative inves-

---

[1] On leave from Dun & Bradstreet, Inc.

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated.................................

tigation of the problems posed would add greatly also to the development of psychology, particularly social psychology.

The available literature, particularly in economics, contains many discussions of the theoretical relationship of psychology and economics (**37, 39, 40, 43, 49, 51, 55, 56, 57, 58, 81, 84, 86, 97, 102, 109, 112, 113, 120, 125, 129, 130, 131, 133, 136, 137, 138, 142, 148, 149, 167a, 174, 175, 190, 195, 199, 200, 202, 204, 212, 212a, 215, 215a, 237, 238, 239, 260, 269, 273, 284, 299**). There are also a number of empirical studies of the psychology of consumer spending and saving, and of labor relations. Few studies of business behavior exist, however, where there is any real merging of psychological and economic concepts and techniques, resulting in new data and insights significant for both disciplines (although see **117, 118, 119, 180, 246**). Moreover, some studies of the economic psychology of consumers (**36a, 36b, 36c, 36d, 119a, 119b, 119c, 148a, 148b, 297, 301**) may suggest principles and techniques applicable to entrepreneurial psychology. This paper will first consider briefly the general relationship of psychology to economics, and will then suggest some of the particular problems of entrepreneurial psychology that, being central to economic theory, merit a high priority in any cooperative psycho-economic research program.

II. Economics as the Study of Behavior in Production and Exchange

Among the economists who have systematically discussed the relationship of psychology to economic theory, and the ways in which psychological concepts and data might contribute to a fuller understanding of economic behavior, there are represented two opposed points of view about the task of economics itself, as a distinctive discipline. One view would define the task of economics as the study of a *particular area of subject matter*, viz., man's activities in production or exchange situations, and would consider of interest to economic theory *all* aspects, or all *measurable* aspects, of behavior in that area. The second view is that economic theory should concern itself only with the *abstract economic aspect* of social behavior, i.e., with the formal implications of the rational pursuit of money.[2]

Economics as an Area of Human Behavior

Up to the early nineteen-thirties, economists of the classical tradition, as well as most of the critics of classical orthodoxy, accepted the *subject-matter* rather than the *aspect* view of the task of economics.

---

[2] This distinction is brought out very clearly by Parsons (**200**), to whom parts of the following discussion are heavily indebted.

Their aim was to develop an economic theory that would make understandable as much as possible of the behavior involved in production and exchange. Whether they started from a priori assumptions about human nature, or from extensive and detailed statistics, both orthodox and unorthodox economists sought to establish general principles that would characterize economic behavior in the "real" world.

Classical Orthodoxy

When these "subject-matter" economists develop economic theory along orthodox lines, they start with five a priori psychological assumptions and, making extensive use of the *caeteris paribus* condition, seek increasing "elegance" in developing the formal implications of these assumptions under this condition (and under additional conditions as they seem pertinent) **(112)**. The psychological assumptions made (although not always explicitly) by orthodox classical theorists are:

1. That production and exchange are predominantly (if not exclusively) motivated by the attempt to *maximize money gains*, usually in the short run.
2. That economic behavior is predominantly (if not exclusively) *rational* and based on reasonably complete *knowledge* of the market and reasonably perfect *expectation* of the results of alternative choices. (Otherwise, assumption [1] would be meaningless.)
3. That additional units of consumption (and hence of income) bring progressively less satisfaction (Gossen's Law of diminishing marginal utility).
4. That total wants inevitably exceed the means of satisfying them (the principle of scarcity).
5. That satisfactions are to some extent substitutable for each other.

Utilizing these assumptions (sometimes called "a priori facts" [**112**]), and requiring that "all other things remain equal," orthodox economists derive general "laws" or "tendencies" (such as the "law" that price will move towards that point where supply and demand are equated, or the "tendency" of the economy to approach general equilibrium). They justify their method by claiming that these "laws" or "tendencies" do, in fact, adequately describe ("explain") by far the largest part of economic (and particularly business) activities.

Critics of Classical Orthodoxy

Orthodox economists do not differ from most unorthodox economists in what they are trying to do (i.e., generalize about a subject matter, a particular area of human behavior). The difference comes in terms of judgments of the adequacy of the method chosen to achieve their common goal. Unorthodox critics of the classical method insist on the importance of additional or alternative psychological assumptions, deny the broad applicability of *caeteribus paribus*, or turn to empirical data (instead of "a priori facts") to provide the bases for their generalizations. Unorthodox economists in this "subject matter" tradition thus seek to build up an economic theory not limited to explaining such economic behavior as is dominated by rational profit-maximizing, but able to explain in addition the great part of concrete economic behavior that is conditioned by social institutions (above all by the forms of business organization), by habits and customs, by population and the laws of its growth, by scientific knowledge and its resultant technology, as well as the great part that is motivated by non-economic considerations, such as satisfaction in doing a job well, in achieving prestige or status, in dealing honestly, in obtaining and exercising power over others, in creating, or in acting in harmony with the "spirit," the value-system, of the particular epoch. When the task of economic theory is defined thus broadly, economic theory tends to become a general theory of social behavior, applied to the particular problems of business and its associated activities. So oriented, economic research contributes to the development of a unified social science whose principles are applicable to any area of behavior.

Adequacy of Psychological Assumptions of Classical Economic Theory

Orthodox classical economists, being satisfied with the adequacy of the conventional psychological assumptions of economics, usually express the best of intentions towards psychology, but do not take it seriously enough to allow it to influence their thinking and research. They consider self-evident the fact that most production and exchange behavior is adequately described by the generalizations of classical economics (modified where necessary to take into account monopolistic or collectivistic organization of production) which result from the logical development of their basic psychological assumptions. Why, then, complicate the structure of economic theory by introducing concepts of behavior motivated otherwise than by profit-maximization, or influenced by non-rational (i.e., impulsive or habitual) factors? Furthermore, to the extent that the economist admits the significance in economic behavior of non-rational or non-profit-motivated factors, he must either measure these factors and introduce them into his theory, or fail in his task of developing valid generalizations about theoretically expected reactions to price and income changes (because he is then faced with multiple equilibria and with indifference curves that cannot be drawn in terms of measurable units.)[3]

---

[3] See, for example, Higgins (108) and Boulding (27).

If economic psychology is considered unimportant by many economists and, even when admitted to be important, is found to be a theoretical obstacle, it is small wonder that psychology gets such summary treatment in most modern texts on economic theory.

And yet, the classical theorist can be quickly led astray by his reliance on the single motive of profit-maximization; for example, consider the classical discussion of interest. Consistent with assumption [1] above, an explicit and more precise psychological assumption underlying the theoretical framework of the famous "stationary state" was that saving (viewed primarily as "abstinence") was largely motivated by the money paid as interest on savings. It followed therefrom that the rate of saving would decline sharply if the interest rate fell below a certain level; hence that "over-saving" could not occur for long, as it would drive the interest rate down until current consumption became more attractive than the command over a slightly larger future consumption.

Much more realistic and, in fact, absolutely essential to an understanding of economic trends was the Keynesian recognition that the rate of saving is *little* influenced by the rate of interest, that saving has a complex and predominantly *non-pecuniary* motivation, and that the act of investment, which implies the loss of the liquidity of accumulated savings, is governed by still another and quite separate group of predominantly non-pecuniary motives (**123**). In this case, at least, the bland assumption of the orthodox classical economists that most economic behavior is dominated by rational profit-seeking leads to palpably false conclusions. This falsity was demonstrated by economists themselves (which may have made the demonstration easier to accept), but it is quite clear that a thorough psychological investigation of the saving behavior of individuals would have brought the classical assumption into question much earlier. It is equally clear that saving behavior should now be studied by psychological techniques, examining into the "arm-chair" hypotheses that Keynes termed "propensity to consume" and "liquidity preference," and developing techniques for quantifying the data on these motives that can be obtained by observation and by interviews with the individuals directly concerned.[4]

There are many other areas where common-sense observation shows the predictive inadequacy of an economic theory that takes into account no motives other than rational self-interest. Workers may turn down employment opportunities at better pay and with better living conditions, preferring to stay in the community and occupation to which they are accustomed, even if that means living on the "dole" in a "depressed area." Businessmen may forego the large profits of smart, aggressive, competitive success for the sake of a quiet, sheltered, habit-drugged business existence so long as their security and position in the industry are not threatened. (See also **151**.) Explanation and prediction of such behavior demands a subtler and more complex approach than classical economic "laws" make possible.

---

[4] The interviews with businessmen carried out by the Oxford economists in the late nineteen-thirties (**7**, **103**, **168**, **224**), and the surveys of consumers liquid asset holdings being carried out for the Federal Reserve Board by the Survey Research Center of the University of Michigan (**119a**, **119b**, **119c**, **148a**, **148b**, **297**, **301**), suggest the kind of research needed to develop a satisfactory theory of interest.

It is important to remember, when considering the "economic laws" of the classical theorist, that these are not "laws" in the sense of generalized descriptions of empirical data. If they were, the confusion about the real nature of the interest rate could never have arisen. Economic "laws" are simply *rational type cases.* They are *norms* for economic behavior, indicating what is to be expected *if* economic behavior is rational and seeks to maximize pecuniary gain.[5] (In addition to the psychological assumptions mentioned above, assumptions are usually made about the degree of mobility of labor and capital in adjusting itself to demand, the degree of competition among buyers and among sellers, divisibility of the commodity, the possibilities of substitution, and the absence of sudden and drastic changes in the economy.) Economic "laws" are *felt* to be empirical, by most economists, because common experience and common sense testify that many people in many situations do in fact seek pecuniary gain and calculate rationally what is to their best advantage. But empirical laws based on scientifically collected data might be quite different from these norms (as in the case of saving behavior), and would certainly be phrased differently. Empirical "laws" would state *how much*, *where*, and *when*, not simply *what.* They would be stated as propositions that could be tested, and retested, against existing or newly collected data (**29, 39, 41, 98, 112, 142, 195, 216, 218, 235, 257, 275, 286**); and they would take into account the great differences between individuals who felt able to influence their economic environment (such as major business or labor leaders) and those who felt they had little such control (small business men, consumers, employees) (Cf. **278a**). One of the first tasks of economic psychology should be to study individuals systematically in their varying roles as consumers, savers, investors, workers, managers, lenders, inventors, landlords, builders, etc., to find out, among other things, how important the rational pursuit of money really is in their economic actions, how its importance changes from role to role and from time to time, and how, in particular, its importance is modified when counterbalanced (or reinforced) by other powerful motives, such as patriotism, need for security, desire for prestige, personal affection, etc.

---

[5] Hutchison (**112**) goes so far as to conclude that the traditional propositions of deductive economics are devoid of any empirical content. See (**113, 120, 137**) for a discussion of this position.

So much for the orthodox classical economists. Satisfied that their "laws" explain most significant economic behavior, especially in the long run, and that economic data obtained by conventional statistical methods are all they need in order to develop and verify those laws, they smile in neighborly fashion at psychology but are careful not to welcome such a disturbing element into their well-ordered household.[6]

To some extent, we have been flaying a dead horse. The critics of classical orthodoxy have been many, and little contemporary economic theory has escaped the influence of their criticisms. However, before considering the unorthodox movements that, in attempting to generalize about the same "subject-matter," have sought to supplement the classical assumptions with additional psychological elements, brief mention will be made of the logically tenable but extreme position vis-a-vis psychology that is taken by the few economists who subscribe to the *aspect* view of the task of economics.

III. ECONOMICS AS THE STUDY OF "ECONOMIZING"

It is only in recent years that the attacks of institutionalists on orthodox economic theory have been met by the highly sophisticated argument that economic theory should not be expected to explain *all* behavior in economic situations, but should limit itself to an analysis of the economic *aspect* of human behavior, leading to the development of abstract economic principles that describe the behavior that *would* result from rational attempts to maximize pecuniary return *if* all other factors (both motives and conditions) were held constant (**52, 97, 112, 133, 149, 174, 199, 200, 215, 215a, 238, 239**).[7] Those taking this position would leave to other abstract bodies of theory (such as psychological theory, sociological theory, political theory, etc.) the task of explaining the rest of the social behavior observed in economic situations. The proponents of this view recognize, of course, that the solution of concrete economic problems and the development of specific economic policies require that economists obtain the cooperation of other social scientists. They point out, however, that scientific progress is possible only through the process of abstraction, and that the time for interdisciplinary cooperation is not at the stage of formulation and theoretical development of general principles, when it merely clouds the issues, but rather at the stage of their application to concrete problems and policies, when it is of transcendent importance (even though its importance is not always recognized.)

---

[6] Weisskopf (278) has made an entertaining analysis of economic theorizing, and especially its choice of certain assumptions and its abstraction from the real world of conflicting interests and anxiety-inducing change, in terms of psychoanalytic defense mechanisms: "isolation" (theoretical abstraction), "projection" (utilitarianism, where man's economic wants "naturally" lead to the "good" of society), "repression" (equilibrium assumed, and antithetical elements repressed), and "compensation" (assumptions of omniscience and perfect rationality made as compensation [fantasy gratification] for some of our main failures). His thesis (278a) that classical economic theory is a projection into nature and into the universe of the money-motivated, unintegrated, and anonymous role forced upon the individual by the machine process suggests that a more rounded economic theory is not likely to be widely accepted except in a society in which the individual has a greater opportunity to express himself *as a totality*.

For these abstractionists, the relationship of psychology to economic theory is very simply stated, as follows. However helpful psychology may be when one is attempting to solve concrete economic problems and frame economic policies, economic theory as such need have no truck with psychology. True, abstract economic theory cannot be considered significant unless one accepts the validity of its basic psychological assumptions (above, p. 291); but common experience can testify to the essential validity of these assumptions (without denying that other psychological factors affect economic action). Furthermore, these psychological assumptions are held to be *constants* so far as economic theory is concerned. Hence psychological measurement of and generalizations about their variations are necessarily excluded.

Orthodox classical economists, subscribing to the subject-matter view of the task of economics, and abstractionists, holding to the aspect view, come out at about the same place with respect to psychology, though for different reasons. The former consider psychology unimportant or a nuisance; the latter consider it irrelevant. Both groups run a double risk: (1) major areas of economic behavior (such as saving or working) may in fact be so dominated by motives other than the rational pursuit of money that theorists must either reach faulty conclusions concerning them or must declare them outside the range of their legitimate interest; and (2) conclusions reached on the basis solely of abstract economic analysis may be mistaken for final conclusions, without taking back into account the psychological (and other) elements that were put to one side for the purpose of simplifying the problem. This "fallacy of misplaced concreteness" (ignoring the incompleteness of principles derived by abstraction) is a fallacy that frequently tempts theoretical economists when they consider concrete problems of policy.

---

[7] A related view holds that economic "laws" simply describe the *conditions necessary* for the stability and autonomy of the market process as a whole. If such laws are inoperative, market equilibrium is impossible; and, in fact, deviation from these conditions may be used to appraise the efficiency of the economic system in getting the most out of its resources. Explanation and prediction of exchange transactions and market processes require the supplementation of these "laws" by the empirical data and principles of social psychology (**150, 151**).

IV. Additional Psychological Factors Introduced into Economic Theory

Economic theory has not, of course, succeeded in remaining "pure." Counterbalancing the tendencies to simplification and abstraction noted above have been strong pressures to recognize the importance of psychological, sociological, political, and legal factors. Considering only the psychological elements, we find that Adam Smith (**236**) recognized that man's "moral sentiments" and regard for public opinion play an important part in limiting the behavior of the economic individualist, and Ricardo spoke of the "habits and customs of the people" as being significant determinants of much economic behavior (**214**, p. 55). It was left to Malthus (**160**), however, to become the first economic theorist really to incorporate psychological elements into equilibrium theory. According to Malthus, economic tendencies could not be analyzed by the assumption of rational profit-seeking alone. One had to take also into account man's procreative urge, leading to a continual pressure of population against the ceiling on consumption set by a niggardly Mother Nature (in the form of the principle of Diminishing Returns).

The Malthusian Theory of Population was for a long time an integral psychological part of classical economic theory. It was eliminated only when empirical data—statistical data—proved it to be invalid, at least for much of the Western World. Its importance as an economic factor in other areas should not be ignored on this account, however. There is still need for detailed studies of the psychological factors influencing the birth rate and the size of families in many parts of the world today, if economic theory is to have general significance.

Psychological Hedonism

Another psychological element became attached to classical economics for a while, without really influencing its theoretical superstructure. Impelled by a felt need to find a psychological foundation for economics, to explain the nature of consumer demand by tying it to something more fundamental, the economists of the nineteenth century, such as Jevons (**114**) and Pantaleoni (**194**), accepted the postulate of a general all-pervading psychological motive, the pursuit of pleasure (and avoidance of pain), and "explained" the demand for a certain commodity by its utility in giving pleasure.[8] "Psychological hedonism" played a part in economic theory for a long time, and turns up occasionally even today.

"Psychological hedonism" did not add anything to classical economic theory, however, for it simply begged the question it set out to answer. The only way one could tell what gave pleasure (of the general sort postulated by the psychological hedonists) was to find out what was demanded (**273**). Furthermore, as was pointed out later, economic theory can be built up quite satisfactorily without knowing anything about what underlies peoples' wants, so long as one knows and can somehow measure those wants (as by knowing how much money will be paid to satisfy them, or which products will be preferred over others) (**43, 53, 106, 107, 130, 138, 186, 221, 222a**, esp. Pt. I; **225, 246, 254, 255, 275, 281**). This does not imply that the area of consumer demand presents no psychological problems to be investigated. On the contrary, we know that the attitudes and motives underlying demand are multiform, unevenly distributed in the population, imperfectly correlated with the ownership of the means for making them effective, variable over time and under the barrage of modern advertising techniques, and intercorrelated in varying patterns; and investigating these problems with adequate social science techniques yields much which is directly relevant to the understanding of economic behavior. But this implies measurement and the collection of new data; "psychological hedonism" was a sterile concept because it did not lead to independent measurements useful in the prediction and control of economic behavior.

The psychological hedonists did not concern themselves with demand only. They wanted also to explain differences in wage rates, and they were not blind to the fact that man's willingness to do certain jobs is greater than his willingness to do others. Taking a rather dim view of work in general, they explained relative willingness to work in terms of the "pain" associated with particular occupations. They did not analyze the matter further.

More acute observers, however, recognized that there are positive as well as negative values associated with work. Marshall analyzed a number of activities that were satisfying in themselves and not as means to ulterior ends, viz., energy, initiative, enterprise, rationality, industry, frugality, and honorable dealing. T. N. Carver introduced into his economic theory the motive of devotion to work for its own sake, his "work-bench philosophy." Veblen spoke of an instinct, or "bent" of workmanship. "Psychic income" is a familiar modern term for the satisfaction derived from the nature of one's work. All these concepts could supplement classical theory, without necessarily denying the domination of profit-maximization in much economic behavior. They all suggested fruitful areas of psychological research, although it was some years before psychological techniques were equal to the task of exploring them.

---

[8] Acceptance of this general principle stimulated the development of elaborate classifications of wants and motives (**22, 22a, 56, 104**, esp. Ch. II; **114**).

Instincts in Economics

Despite the sterility of the concept of psychological hedonism, the desire to base economics on a psychological foundation continued strong, and economists somewhat slavishly followed the next fashion in psychology. The chapter on instinct in James' *Principles of Psychology* (published in 1890) may be said to have started the conversion of economists from complete dependence on the rationalistic hedonism of the utilitarians. Following the publication in 1908 of McDougall's famous "Social Psychology," the instinct theory was adopted by many economists, impressed to find the vagaries of human behavior so easily and clearly "explained." They promptly started writing economics in these new psychological terms (54, 55, 56, 61, 80, 110, 161, 172, 176, 196, 197, 237, 247, 249, 250, 251, 268).

Despite its popularity, the instinct hypothesis did not represent much of an advance in social science. Experimental investigation and systematic observation (primarily by the behaviorists) quickly showed that humans did not inherit rigid patterns of response to social situations. Every writer had his own list of instincts, comprising groups of activities in which he saw some unifying element; but there was no way to identify behavior motivated by the "instinct of pugnacity," for example, except to agree that certain behavior appeared pugnacious. The terms "pugnacious instinct," "instinct of accumulation," etc., could not mean anything operationally for prediction and control until measurements of the instincts could be carried out independently of the economic behavior they were invoked to explain; and psychologists were unable to provide such independent measurements.

On two scores, however, the instinct fad was a step forward. The variety of human motivation was stressed (in contrast to the generalized pleasure-pain formula of the psychological hedonists), opening the way to research into particular attitudes and motives; and the analogy to animal behavior drew attention to the need for objective investigation, in contrast to the introspective, arm-chair approach.

In another direction also, the instinct theorists deeply influenced economic thinking. Their work cast serious doubt upon the rationality of economic behavior. Psychoanalytical writers were also emphasizing the importance of non-rational, subconscious drives, and the unreliability of "rationalizations;" while behaviorists were, on quite different grounds, beginning to argue that social scientists should disregard all introspective data. Economists consequently found it difficult to continue to think of conscious, rational profit-maximization as the dominant guide to economic action. This had two unfortunate results. On the one hand, interest in classical economic theory diminished and its development was retarded. On the other hand, those who continued theorizing in the classical tradition cut themselves off from contemporary currents of psychological thought, and their distaste for and distrust of psychology in economics endures to this day.

Veblen occupied a unique position, in this respect. He was an instinctivist, postulating three ultimate instincts pertinent to economic action (parental bent, bent of workmanship, idle curiosity); but his instincts were explicit, conscious "ends of action." Pecuniary reward was not, however, one of these ultimate goals, and he denied that the rational pursuit of money was an organizing principle in economic society. Moreover, he and Marx were among the few economists willing to recognize that goods and services were very widely obtained *outside* the market, in ways other than by exchange (i.e., by gift or by seizure) (see **84**, pp. 11–15; **250**; **272**, ch. 6). Unlike the critics of classical economics mentioned above, Veblen wanted not merely to broaden existing theory; he wanted to replace it with a theory he found more satisfying (**6**, **52**, **84**, **263**, **264**, **265**, **266**, **267**, **268**, **269**).

Farthest away from the classical tradition was the economic theory of the German Historical School, of which Marx and Sombart were the most prominent members. According to them, there was no possibility of a universally valid economic science, because the ultimate goals of economic action, the motives underlying the organization of production, were impermanent, were not characteristic of man in general, and were applicable only to a single epoch in man's history. (See also **149**, **277**, **278a**.) Thus, the study of the psychological motives underlying economic behavior was not simply a method of broadening one's economic theory; it was essential to an understanding of that theory, and was the only way one could tell whether that theory had any validity for the present day and age.[9]

---

[9] Classical economists themselves were not as thoroughly committed to the concept of a universally valid economic science as sometimes represented. J. S. Mill pointed out that a person is not likely to be a good economist who is nothing else: ". . . . social phenomena acting and reacting on one another, they cannot be rightly understood apart. . . . . the material and industrial phenomena of society are . . . . susceptible of useful generalizations, but . . . . *these generalizations must necessarily be relative to a given form of civilization and a given stage of social advancement.*" (**171**, p. 81) (Italics mine.) There are also suggestions in Marshall (**165**) that ultimate wants are not fixed, that they may become "adjusted to activities," changing in nature as the organization or technology of society changes (see Parsons, **198**). Clark (**40**) has an interesting discussion of the possibility and economic implications of the malleability of wants.

Institutionalism and Later Unorthodox Developments

Since Veblen's day, most efforts to add to the parameters of economic theory have been grouped together under the broad and rather amorphous concept of "institutionalism," meaning in general (1) dissatisfaction with the adequacy of classical theory, (2) the suggestion of additional principles governing economic behavior, and (3) an emphasis on the importance of objective facts, versus deduction from logical but not necessarily empirical principles (**6, 37, 49, 51, 52, 56, 57, 84, 89b, 93, 109, 129, 131, 167a, 173, 174, 175, 190, 191, 195, 200, 201, 203, 208, 209, 212, 212a, 212b, 228, 237, 250, 260, 261, 282, 283, 284**). Institutionalism has done much more to disrupt theory, however, than to improve it. Without formulating or making possible the formulation of a new theoretical structure, institutionalists attack the adequacy of the old, thus becoming in effect not merely anti-classical, but actually antitheoretical.

The failure of institutionalists to erect a new theoretical structure may be ascribed to something more fundamental than an antitheoretical bias. To build, one must have tools, and the most essential tools (for this purpose) are measuring instruments. For the classical theoretical structure, money provides the measuring instrument. Demand (though not all demand) can be measured in money; costs (though not all costs) can be measured in money; and reactions to economic changes can be predicted by assuming the dominance of money-profit-maximization. As Marshall emphasized (**165**, Bk I, Ch. II, and Appendix D, par. 2), the theoretical structure of economics could not have been elaborated without the measuring instrument provided by money, for the progress of science depends on measurement, and other measuring instruments have been wanting.

What the institutionalists lacked, and what prevented their development of a new theoretical structure, was measuring instruments for the determinants of economic behavior that are not expressible in monetary terms, that is, for most motives, attitudes, habits, expectations, etc. Fortunately, this lack is today being supplied—at least in part—by the new techniques of social science. The application of these new techniques will be discussed later in this paper. Here, it is enough to point out that we are on the verge of a new flowering of economic theory, a theory that will take into account (in measurable terms) many of the psychological elements whose exclusion from economic theory has so long been protested by the critics of classical orthodoxy.

Related to the institutionalist tradition are two recent developments in economic theory, viz., (1) the emphasis on psychological factors in dynamic process analysis, and (2) the controversy over the adequacy of marginalism in the analysis of the firm. It is interesting that, with the shift in interest in macro-economic theory from emphasis on static equilibrium analysis to an emphasis on dynamic process analysis, there has come also an increased recognition among economists of the importance of psychological elements in economic behavior. For example, of the four factors stated by Keynes (**123**) to be adequate to explain the level of employment, three are psychological, namely "the propensity to consume," "liquidity preference," and the subjectively expected "marginal efficiency of capital." In most dynamic process analysis, "expectations" play a key role.[9a] Econometricians usually leave room for psychological factors to be introduced into their mathematical models, if and when psychologists can measure such factors. Most other economists concerned with public policy likewise stress the significance of psychological factors in determining the level of economic activity (e. g., **183**, esp. pp. 49–51; **188**, esp. p. 27; **234**, esp. pp. 43–62).

In micro-economics, also, the importance of supplementing pecuniary marginalism with other elements has gained increasing support. Following the publication of data obtained by the unorthodox but sensible (although often technically unsophisticated) method of interviewing individual businessmen (**7**, **8**, **24**, **67**, **92**, **103**, **117**, **144**, **158**, **168**, **223**, **224**, **294**), an active controversy has raged concerning the adequacy of the marginalist approach to the theory of the firm (**48**, **53**, **69**, **89**, **111**, **144**, **145**, **146**, **146a**, **156**, **157**, **192**, **193**, **211**, **213a**, **245**). In summing up the evidence, R. A. Gordon (**89**, pp. 287–8) concludes that traditional marginalism has been pretty well worked out, that further empirical research limited to the tools of marginalist theory is likely to be sterile, and that significant advances in economics are now to be expected only if empirical workers are supplied with new and more useful analytical tools. (Also see **26**, pp. 801–802; **97**, **142**, **167a**, **216**, **218**, **272**, **286**.)

As we survey the history of economic thought, it is apparent that economists, even in their theories, have never been satisfied with the concept of the purely "economic man." They have had to consider many of man's behavior traits. As Mitchell points out, "Keynes' concepts of the 'propensity to consume' and 'liquidity preference', so confidently invoked today, are as patently psychological as Adam Smith's 'propensity to truck, barter and exchange', or Bentham's 'felicific calculus', or Malthus' 'instinct of procreation', or Bagehot's 'cake of custom', or the Austrians' 'marginal utility', or Edgeworth's 'indifference curves', or Veblen's 'cultural incidence of the machine process', or Schumpeter's distinction between 'routineers' and 'innovators', or Pigou's epidemics of 'over-optimism' and 'over-pessimism', or Freud's 'complexes' hiding in the 'subconscious'. The 'economic man' was a psychological caricature, deliberately drawn to facilitate speculation, but condemned in advance to be a sketch of how creatures very differently endowed than human beings would behave" (**178**, pp. 16–17).

[9a] Analysis of expectations is just as subject to highly refined deduction from unsubstantiated psychological assumptions as is any other sector of economic theorizing (**139a**, **167a**, **231**, **232**).

The goal of any science is generalization, prediction, and—eventually—the ability to control. For any area of economic behavior, the goal of economic science should be to formulate valid generalizations concerning the conditions under which certain actions take place, to predict that certain conditions will be followed by certain actions, and to prescribe the steps to be taken to bring about the particular actions desired. If people could in fact be depended upon to act "*as if*" profit-maximization were their goal, and rational balancing of alternatives their method, there would be no need to complicate economics with more psychology. If further psychological assumptions promise to improve prediction and control, however, it would be absurd to exclude psychology from economics on the grounds either that it seems a priori to be unimportant, or that its inclusion would destroy the distinctiveness of the economic discipline. Most economists undoubtedly consider that they have a legitimate interest, as economists, in *all* of the factors influencing the organization and processes of production and exchange.[9b] The crux of the matter is whether or not psychological data, principles, and research techniques promise to increase substantially our capacity to understand, predict and control economic behavior, and this question can only be answered after the fact, not before it.

V. THE PSYCHOLOGICAL APPROACH TO THE STUDY OF ECONOMIC BEHAVIOR

In general, the psychological approach in studying behavior differs in important ways from the economic or the statistical approach. First, data are sought concerning the behavior of *individuals*. Aggregates, averages, and indexes may be utilized, but only after the *individual* variations underlying them have been investigated. Second, data are sought by going *directly* to the individual, not relying on statistical records through which behavior can only be studied after the fact and at second- or third-hand. Third, *all* kinds of habits, attitudes, and motivations are considered, not simply profit-seeking. And finally, consideration is given to the ways in which attitudes and motivations *change*, and the techniques that might be useful in changing them. Investigators other than psychologists may approach economic problems in these ways, but these are the usual characteristics of the psychologist's approach. (Cf. **34**, **102**, **119**, **126**, **148**, **212**, **212a**, **212b**, **237**).

[9b] Even those economists whose interests lie primarily in the development of formal theory along classical orthodox lines recognize the importance of empirical research in cooperation with other social scientists (**64**; **215**, ch. IV; also see **39**, **58**, **81**, **109**, **149**, **167a**, **272**, **273**, **284**, **285**, **299**).

Emphasis on Individual Differences

Psychologists do not have a monopoly on the individual approach. The whole section of economic theory that has to do with the behavior of the firm is essentially a study of hypothetical (and sometimes real) individuals (see above, p. 302). Moreover, the realization of the actual (and growing) extent of imperfect competition and the resulting possibility of administered prices has led to increasing emphasis on understanding the behavior of the individual oligopolist, whether price leader or follower.

Quite as important as understanding the individual entrepreneur, of course, is understanding other strategic individuals in the economic arena—the labor leader, the policy officers of great financial institutions, government officials, leaders of opinion and fashion in the community—whose economic decisions have an impact far out of proportion to their number. Today, we live in what is to a large extent an *administered* society, rather than a society whose character results from the interplay of countless decisions, among which none has a disproportionate effect (**86**, **105**, **185**). An understanding of the bases of administrative decisions is now a necessary supplement to, and perhaps even outweighs in importance, an understanding of the statistical tendencies of mass opinion and action.

Obtaining Data at First Hand

In studying individual behavior, the economist has generally been content to use as data the records collected by others—records of new orders, posted prices, plant construction, labor turnover, etc. These data may be unreliable or unsuited to his needs (because collected with other needs in mind). They are often poorly dated and nearly always too superficial to reveal what really took place at the time the essential economic decisions were made (**34**, pp. 6–8; **102**, pp. 87–89). The use of psychological research techniques (such as the case study,[10] sample surveys utilizing depth interviews, content analysis, or experimental social psychology) could be expected to illuminate the behavior *underlying* these surface records.

Each individual who might be the subject of psycho-economic study typically plays more than one economic role.[11] He is certainly a consumer, probably also an employee or entrepreneur, probably also a saver, possibly also an investor in securities. Other possibly overlapping roles are those of the government policy-making official and the civil servant, the financier, the employed manager, the engineer or other professional staff man, the voter, etc. While the same individual may and usually does play more than one of these roles concurrently, he usually acts in any given economic situation as though he were playing only one role. For the time being he is dominated by a conception of himself as a particular bundle of habits, attitudes, objectives, etc., and he may and often does act in one role in ways that are inconsistent with his behavior in other roles (e.g., an entrepreneur or an employee may seek a higher direct return from an enterprise even though the result may be that as a consumer he has to pay higher prices). There are very great differences in the groups of motives dominating the different economic roles man plays, and there are great differences in the motives dominating the different kinds of economic action that occur within any one role. For meaningful psycho-economic research, therefore, it is necessary to break down the variety of economic behavior into sub-classes having some homogeneity—both in terms of economic significance and in terms of superficial psychological similarity. In section VI, there is suggested an a priori breakdown of classes of action taken by individuals playing the role of *entrepreneur*. This breakdown indicates the great variety of separate psycho-economic investigations needed for each role played by economic actors.

---

[10] A particularly promising technique in studying business behavior is the case-study. If mature investigators, well-trained in economics, social psychology, and interviewing techniques, could sit in regularly on policy discussions of industrial concerns, and could interview intensively the significant policy-makers over extended periods of time, they should be able to develop a body of case material showing *how* policy decisions are arrived at, what considerations are given what weight, etc., that would add greatly to our present knowledge and understanding. This technique is now being utilized in the Brookings Institution study of industrial concentration and in the University of Illinois study of business expectations. Also see **206**, **213a**, esp. pp. 305–308.

Limited Area Theories

Studies that are each limited to a particular class of economic action may be expected to result in the supplementing of general economic theory by a series of particular theories of economic behavior, each comprising generalizations or "laws" that are empirical (rather than "type cases") and each appropriate to a given area of economic activity—these areas perhaps being defined both in economic terms and in geographic terms. For example, farm ownership behavior may be quite different in Pennsylvania, where the descendants of Dutch settlers abhor being in debt for anything, from what it is in the South, where it is traditional to borrow even against current crops. Similarly, the tradition of the textile industry, where one gets ahead by competitive cost and price cutting, contrasts sharply with that of the metal industries, where much more time and effort has gone into attempts to work out *non*-competitive arrangements (cf. **26**, **101**).

---

[11] Economic "role" is not meant to refer simply to one's *economic interests* (as a consumer, entrepreneur, etc.), but to the much more significant concept of the *entire* behavior pattern appropriate to a particular role. This "picture in the mind" of what is *fitting*, which takes into account *all* one's attitudes and motives, provides a much more complete and reliable basis for predicting economic behavior than does the limited concept of rational economic interest.

As a separate group of psycho-economic generalizations (a "theory") is developed for each area of economic activity, it will become possible (1) to predict reactions to purely economic changes and controls; (2) to develop a range of possible supplementary psychological controls; and (3) to predict the likely success of these psychological controls. The prediction and control techniques developed may be expected to be tailored differently for each class of economic actions found to be psychologically distinct, if the greatest effectiveness is to be attained.

Variety and Malleability of Motivations Affecting Economic Behavior

A major difference between the approach to economic research dictated by classical (including most current) economic theory and the approach utilized by the psychologist lies in the variety of determinants of economic behavior considered significant, and in the constancy attributed to them. The economist's theoretical interests are limited to a few psychological determinants (those that impinge, or are believed to impinge, on profit-maximization) and he tends to look upon these as relatively unchanging. His theory has been unable to shield him from the facts, however, and the distinction between the economic approach and the psychological approach is consequently fuzzy. The empirical literature of economics is studded with insights into the non-rational and non-pecuniary determinants of economic behavior, and with gropings for psychological principles that can supplement the principle of rational self-interest. These tentatives are usually isolated, unsystematized, and undeveloped in terms of their psychological origin, function and characteristics. Nevertheless, they demonstrate how strongly economists have felt the need to widen and deepen the psychological foundation of their theoretical and predictive structure. This economic psychologizing constitutes a standing invitation and a challenge to psycho-economic investigation, for its significance to a fuller understanding of the economy is quite apparent from the importance of the problems which it is called upon to help solve.

Two fields of psychology are consequently of particular significance to economics, viz., *learning* (including communication) and *motivation*.

The principles of habit-formation, the extent of habitual behavior in economic situations, the significance of the "Gestalt," the characteristics and incidence of insightful learning, and the importance of suggestion, imitation and sympathy, have many implications for the understanding, prediction and control of economic behavior (**118**). Advertising and propaganda are based upon them. Much economic behavior can only be understood in the light of the advertising and propaganda (and sometimes the education) that have impinged upon particular economic groups. And the achievement of control of economic behavior (whether by private groups or by government) often rests more upon adequate use of techniques of communication than upon economic or legislative actions. Human beings are malleable, and economists who take man's economic psychology as a given, without caring how he got that way or what possibility there is of changing him, are contenting themselves with a much less effective role than they might play.

In the field of motivation, also, there are psychological concepts and principles of particular value to economists. There is special significance in the modern understanding that motives are not fixed and invariant, or universally present (except in very general terms); and that wants are all cultural, at least in the forms in which they are significant.[12] Most needs, wants, desires that play an important part in economic behavior are oriented toward fairly specific goals. They have, however, become attached to those goals during the individual's prior experience, and, within limits, they can be detached from one goal and attached to another. Within broad classes, there is thus the possibility of substitution of one goal for another. But this malleability and substitutability do not mean that the existence of presently effective motives cannot be demonstrated, nor that the incidence and relative strength of these motives cannot be measured. To a degree, these motives are measured indirectly by market behavior, choice of vocation, volume of saving, etc.; but direct investigation of particular individuals can be a much more fruitful method of identifying and evaluating the strength of these motives, in such a way that economic behavior can better be understood, predicted and controlled. No investigation of particular economic behavior is complete unless the investigator discovers:

1. What persons are active and influential in making significant economic decisions? How can these policy and decision makers be identified, and what techniques of investigating them promise to be most fruitful? What kinds of interaction are involved in the process of decision-making?
2. What attitudes, beliefs, expectations, habits, knowledge, and motives are important, and how influential are they in each decision-making situation? How many (and which) people have these attitudes, knowledge, etc., and how strongly? Which attitudes are reliably correlated with *action?*
3. What is the origin and developmental history of these attitudes, etc.? What are their sources in the culture (general and specific culture areas)? How are they transmitted (through what educational channels, what media of communication, what kinds of opinion leaders, etc.)?
4. How malleable are existing attitudes, how receptive are economic actors to new knowledge, etc.? How does receptivity to new knowledge, rumor, sug-

---

[12] Many economists, particularly among the institutionalists, have stressed this concept, but all have been content with descriptive accounts (**4**; **5**, Ch. I; **21**, **40**, **47**, Pt. VI; **56**, **125**, **136**, **165**, Bk. III, Ch. II; **190**, Part I; **198**, **202**, **204**, **210**, **264**).

gestion, etc., vary among individuals, and what factors are related to such variations? What psychological techniques are available to influence attitudes and through them each class of economic actions? How effective can these techniques be made?

Research Needed Even with the Simplifying Assumption of the "Economic Man"

Although most economists (like other social scientists) recognize the inadequacy of the single motive of "rational self-interest" as a determinant of economic decisions in real life, there is a wide area where psycho-economic research is needed even within the limits of this concept of motivation. Every economic decision is made in a particular economic *situation*, comprising (1) all the elements of the institutional framework (laws, political system, customs, etc.), (2) all the economic data (levels of wages and prices, production costs and sales volume, level of profits and interest rates, strength and efficiency of competitors, etc.) *known* to the actor and believed by him to be *pertinent* to his decisions, and (3) all of the actor's *expectations* about the future behavior of these institutional and economic factors.

Knowledge of pertinent facts not a safe assumption. There is first the question of what facts and factors are considered by the actor to be *pertinent* to his decisions. There has been a great deal of arm-chair speculating about the kinds of facts and factors that *should* be recognized to be pertinent, if the economic actor were analytical and farseeing. (Appendix A outlines many of these institutional and economic factors.) There is very little that is really known, however, about the range even of the strictly economic considerations that influence economic behavior, and even less that is known about their relative importance as influences. Furthermore, this is a field where belief is as important, and probably more widespread, than knowledge. The gyrations of the stock-market are *believed* by many businessmen to be significant for them, not because of their economic effects (on the interest rate, on the ease of financing, or on their own personal finances) but because of their presumed psychological effects on the expectations of other business men. Most businessmen *believe* that certain economic statistics are particularly significant as indicators of changes in business conditions, but the actual significance of most such indicators is obscure and considered by business-cycle students to be meaningful only in the context of the whole economic situation. Economists have often stressed the importance of the interest rate, and assumed that it would be an important element in business calculations; yet few business men apparently consider it very significant for their own operations (7, 67, **103, 168, 224**).

Again, economists speculate about the theoretical effect on economic behavior of the absence of complete knowledge of pertinent data, but there is practically nothing known about the actual *extent* of knowledge among various economic actors today. How much do entrepreneurs, consumers, employees *know* about the facts of their own economic situations (e.g., **26**, **81**, pp. 61–63; **153**)? What are the sources of their information, and how reliable are these sources? How can their knowledge be increased? What changes in behavior ordinarily result from more knowledge of particular kinds? Here is a whole world that needs exploration.[13]

Expectations a key field for psycho-economic research. Even more important for economic decisions than how much is actually known about current economic conditions, however, are *expectations* about future economic trends, about competitor reactions, about governmental taxation and controls, etc. The *concept* of expectations is central in modern economic theory (**99, 100, 106, 112, 118, 128, 139a, 180, 226, 231, 232, 256, 271**). The *facts* about expectations are almost unknown and constitute a major field demanding psychological investigation. Because of the importance of expectations for all the fields of economic behavior, some of the questions about expectations that should receive early attention in a psycho-economic research program are outlined here:

1. *Whose* beliefs and expectations are important? How identify the economic opinion leaders? Flow of opinion from whom to whom, and through what channels (**17, 17a**)? Are industries whose turning points "lead" the business cycle headed by business opinion leaders? Note that expectations of large companies and of "price-leaders" are much more important than those of small; sampling should stress those individuals who change *first*, and who have the greatest impact.

2. How do expectations develop (**117, 226**)? What media containing economic information, business opinions, and forecasts reach what people (**205**)? What are taken how seriously? What factors give recipients confidence in the reliability of their informants? How important are wish-thinking and differences in susceptibility to rumors of different sorts? Can the effects of long-run susceptibility be separated from those of short-run predispositions resulting from temporary insecurity?

3. How easily do expectations change? Under what conditions is change facilitated? What psychological techniques may be effective in changing business expectations (**18, 298c**).

4. How strongly are expectations held? Can strength of conviction (resulting in action?) be measured separately from persistence (unwillingness to consider contrary information)? Are expectations typically much more confidently held and more slowly changed when they refer to consumer behavior than when they refer to business (and particularly *investment*) behavior? Is certainty or uncertainty greater at the height of a boom?

---

[13] The fields of market research, commodity research and cost accounting have, of course, been very greatly developed in the last twenty years to meet this very need. Little is known about how widely these kinds of information are utilized, however, it is important, furthermore, to recognize the great significance of particular statistical methods in influencing business behavior. Different accounting assumptions, for example, may introduce important biases into what is "known" (**77, 241, 294**).

5. How variable are expectations at any one time? Can variance be used as a predictive factor?

6. Is there greater unanimity about "general business" than about one's own industry or concern? Are there several "climates of opinion" held by the same people? Which affects business behavior most importantly, and how (**117, 291, 300**)?

7. Methods of measuring expectations of business men (**117, 180, 253, 291, 292, 300**) and of farmers (**31, 72**, pp. 428–429; **296, 298b**).

8. Measures of validity of expectation data, including reinterviews to determine how subsequent action is related to expectations, and correlation of expectation series with business activity series (**33**). Is there a tendency for expectations to "justify" themselves (**18, 28**, p. 138)?

9. Nature and influence of individual concern's expected demand curve (**32, 139a, 213a, 231, 232, 244, 259**, esp. pp. 78–108; **271, 298b**).

10. Nature and influence of expectations about future course of wage rates, materials costs, selling prices, etc. (**20, 106, 117, 139a, 184, 287**).

11. Changes in business motives in the light of changing expectations (e.g., high profits sought during business expansion, security sought during contraction) (**1**).

12. Measurement of impact of particular events—such as changed government policies—on business "confidence." Importance of knowledge that particular monetary, fiscal or other governmental measures will be taken in stated circumstances (**163**).

13. Effects of changing business expectations on price and production decisions (**78, 117, 139a, 231, 232, 298b**).

Beyond the "Economic"

It is obvious, then, that, even accepting the simplifying assumption of a single economic motive of rational self-interest, there is plenty of need for psycho-economic research; but the field is far wider. Most economic behavior in real life also involves other motives and much behavior of an habitual character. From a psychological point of view, economic behavior is as complex as any other social behavior, and requires as comprehensive an approach—with only the limit that the behavior (decisions) studied should have generality and be of significance for the economy.

Within economics as a whole, the behavior of entrepreneurs (including their investing behavior) has received the greatest attention in economic theory, and at the same time has drawn the least on psychological concepts and techniques. Neither the behavior of workers nor the behavior of consumers has received nearly as much attention from theorists, and both have had much more attention from psychologists.[14] The outline below is limited to the behavior of entrepreneurs, and to the areas of entrepreneurial behavior that are particularly promising for inter-disciplinary research. It is the purpose of this outline (1) to point out the areas of entrepreneurial decisions that have special significance for economic theory, (2) to indicate the variety of nonpecuniary motives (and some of the psychological complexity of the "basic" profit motive) to which economists themselves have found it necessary to appeal in developing generalizations about economic behavior, and (3) to aid prospective psycho-economic teams to find the literature (primarily economic) that will be helpful in setting up the theoretical economic questions that need empirical investigation by such teams. (These references are not exhaustive, however, nor do they constitute the source of all the psychological factors listed.) It is the hope that presentation of this outline will stimulate psycho-economic research that results in the quantification of the incidence, influence, origin and malleability of the factors listed.

---

[14] Theoretical studies of the consumption function are, of course, important and fairly numerous, and these are increasingly tied to empirical investigations—some of which draw heavily on the psychological approach.

VI. Critical Areas of Entrepreneurial Behavior Requiring Psycho-economic Investigation

Meaningful investigations of enterpreneurial behavior can only be designed if the major classes of entrepreneurs are distinguished and studied separately, thus holding constant some of the important variables related to economic behavior (**112**, esp. pp. 163–165). It is apparent that the effective psychological factors, and hence the likely decisions, are quite different for individual farmers from what they are for corporation farmers, or for small retail trade and service entrepreneurs, for large corporations engaged in retail and wholesale trade and services, for small or large manufacturing concerns, for small or large banking and financial establishments, for publicly regulated utilities, or for government owned and operated enterprises. The economic populations selected for study should, besides being categorized as above, be further sub-classified in such a way as to eliminate the effects of variations in the characteristics of economic activity and institutions outlined in Appendix A.

As entrepreneurs carry out their economic functions, the areas of decision most important for economic theory are: entering or leaving an industry, keeping one's capital more or less liquid, raising more capital in different ways and from different sources, introducing technological or organizational changes, formulating price and production policy, gaining market advantages through means other than price competition, selecting distribution channels and sources of supply, and determining wage levels. (Decisions concerning procedures and techniques of employment, sales promotion, industrial relations, and participation in political activity are less important for economic theory.) These major areas are treated separately in the following outline.

A-1. *Area of decision to be taken by entrepreneur: whether to enter an industry, continue in it, expand in it, or leave it.*

Decisions to establish an enterprise; to buy into a going enterprise; to increase investment in an enterprise, or the scale of its operations; to take on additional lines, produce additional products, offer additional services; to plough profits back into the business and thus expand investment; to maintain investment, refusing to expand in any way; to make inadequate allowances for depreciation and thus reduce investment; to drop some lines, products or services; to withdraw capital or reduce the scale of operations; to sell a portion of the enterprise; to sell or disperse the whole operation.

A-2. *"Non-economic" factors considered likely to influence decisions* (excluding economic data and expectations).

Extent of personal participation, as where the enterprise becomes one's major occupation or even a *way of life* (as in farming, small retailing, or the professions).

Importance of personal independence.

Prestige or social status of particular industry, or of ownership or managerial position *vs.* being an employee of someone else.

Mechanical, commercial or professional aptitudes and interests.

Scope for imagination and inventiveness.

Desire to participate in constructive enterprise—the feeling of *creating* something.

Desire to build or preserve a family institution and economic base.

Family or cultural traditions.

Readiness to take risks.

Urge to dominate.

Desire for security of capital or stability of income (rather than maximum total profit).

Attitude towards "rights" of stockholders, *vs.* managerial desire to build up concern.

Sense of responsibility to owners, employees, customers or community for maintaining and perhaps building up economic activity.

Readiness to take advice or get training on conduct of business, and on factors necessary for success (vocational counseling, small business institutes).

Adaptability of management, workers and organization to new design or new product; or their attachment to established line.

References: **8, 9, 15, 26, 27, 36, 46, 62, 65, 71, 73, 82, 85, 87, 89a, 100a, 114a, 115, 123, 127a, 146a, 155, 164, 222a** (esp. Pt. I), **227, 237, 248, 262, 266, 268, 279, 280, 293.**

B-1. *Area of decision to be taken by entrepreneur: what form to keep existing capital in.*

Cash, government notes, government bonds, other securities unrelated to conduct of business, securities providing some measure of control over related business operations, accounts receivable (including consumer credit), loans (commercial or personal), inventories, plant, equipment, real estate, patents, technical processes and other research results; training, welfare or morale of labor and management; goodwill of consumers, distributors, suppliers, employees, or general public.

B-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Felt need for "safety-margin" of quick assets.
"Liquidity preference."
Ability to judge credit-worthiness (e.g., "character" appraisals of applicants for credit).
Belief in ability to judge credit-worthiness.
Felt security of capital in each form.
Traditional or habitual financial practices in the industry or company.
Desire to control operations where capital invested.
Desire to reduce or prevent competition with existing interests.
Desire for flexibility provided by liquidity of assets.

References: **10, 45, 70, 73, 74, 75** (esp. p. 152), **85, 90, 95, 96, 99, 100, 100a, 106, 123, 127, 128, 146a, 164, 170, 274, 276, 298a.**

C-1. *Area of decision to be taken by entrepreneur: whether to raise additional capital.*

To strengthen current financial position; to strengthen control (as by owning rather than leasing property); or to enable expansion of operations.

C-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Attitudes toward owned versus rented property.
Attitude towards being in debt.
Desire to control enterprise.
Ability to tolerate insecurity arising from shortage of liquid capital assets.
Desire for flexibility provided by liquidity of assets.
Attitude toward expansion of operations (see A, above).
Belief in importance of repaying debts, in "sacredness" of contract.
Attitude towards going through bankruptcy.

References: **7, 8, 67, 91** (esp. pp. 39–44), **96, 103, 124, 132, 134, 135, 139, 140, 141, 152, 164, 168, 222, 224, 227, 230, 234, 252, 258, 270, 274, 288, 289.**

D-1. *Area of decision to be taken by entrepreneur: how to raise needed capital.*

Loans from banks, RFC, or other lending institutions; issue of bonds or debentures; issue of preferred or common stock; borrowing from officers or other individuals; or holding back earnings for reinvestment.

D-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations)

Attitudes toward desire of creditors for some measure of control over policies and operations.
Attitude toward stockholders' "right" to control reinvestment of earnings.
Effect on prestige of borrowing at a high rate of interest.

Attitude toward open market treatment of securities, with their implications for prestige of company.

Insecurity generated by need to make regular payments (e.g. on loans).

Reference: **274.**

E-1. *Area of decision to be taken by entrepreneur: whether to change technology or organizational arrangements.*

New or improved tools or machinery, or strains of plants and animals (in farming); work flow and factory layout; activities or functions based on time and motion studies; supervisory responsibilities; customer participation (self-service or self-assembly); utilization of new types of professional or managerial advice; statistical quality control; operations analysis and market research; cost analysis and accounting methods; encouragement of research; suppression of inventions.

E-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Receptivity to innovation on part of management or labor force.

Relations between management and labor force.

Anticipated competitive position.

Anticipated control over improved technique.

Stage in growth of an industry (new techniques easier to introduce during a period of expansion).

Anticipated effect on employment, wages, conditions of work.

Attitude towards discarding usable but obsolete equipment.

Attitude towards going into debt.

References: **19**, **25**, ch. 9; **44**, **66**, **79** (pp. 70–80), **121** (n. 4, p. 170), **159**, **162**, **169**, **181**, **182**, **242** (esp. pp. 59–66), **295**.

F-1. *Area of decision to be taken by entrepreneur: price and production policies in the short-run* (i.e., involving no change in investment).

Expanding or restricting volume of production; altering prices directly or indirectly (i.e., by changing quality, services provided, terms of sale, liberal credit, etc.); changing selling costs, including relative expenditures for advertising, public relations programs, etc.

F-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Beliefs about what industries are really competitive with own.

Expected reaction of sales to price change. (Expected elasticity of demand.)

Relative importance of "lucrativity, safety or asymmetry" of expected net receipts.

Extent of understanding of marginalist concepts and thinking in marginalist terms.

Belief in a "fair" price or a "fair" profit.

Belief in high-volume-low-profit policy vs. high-profit-restricted-volume policy.

Group attitudes of producers toward "price-cutting."

Identification with, and importance of prestige and status among, producers group.

Community attitudes towards charging "fair" prices.

Belief in importance of goodwill of consumers, distributors, employees or public at large.

Identification with, and importance of prestige in, community.

Identification of own interest with that of customers.

Expectations of retaliation by customers when conditions change or even in other relationships (ostracism, voting of governmental controls, even physical violence).

Feeling of obligation to ownership interests.

Governmental controls, and attitudes of group, community and individual towards compliance therewith.

Desire for security, liquidity, or flexibility of capital or for maintenance intact of going organization.

Desire to demonstrate personal independence or power.

Belief in short-run and long-run effectiveness of advertising and public relations programs.

Willingness to seek and use advice on management policy.

Willingness to adapt policies and management to competitive pressures.

Belief in and willingness to use particular accounting practices (such as average vs. marginal unit costs, proper differentiation of capital expenditures and operating costs, different depreciation policies, etc.).

Closeness of personal relationship between decision-maker and purchaser.

References: **6, 7, 8, 14, 14a, 23, 26, 29, 31, 45, 53, 76, 77, 78, 83, 89, 89a, 92, 94, 98, 102, 103, 108, 111, 116, 117, 123, 127a, 143, 144, 145, 146, 146a, 158, 166, 168, 177, 179, 187, 189, 207** (esp. chs. VI & VII), **211, 220, 222a** (esp. Pt. I), **223, 224, 226, 229, 237, 240, 243, 279, 290.**

G-1. *Area of decision to be taken by entrepreneur: how to maintain or strengthen market position by means other than price competition.*

Advertising; differentiating products; making production and marketing agreements; licensing agreements; basing point systems; mergers; trade associations; obtaining legal limitation on entry of new concerns, on level of production, on increased investment, on sale in particular areas, or on price-cutting ("fair-trade" laws), etc.; cooperative marketing; cooperative purchasing; cooperative production; patents; or use of non-economic pressures, such as friendship, persuasion, ostracism, publicity, etc.

G-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Feeling that competitive results are too harsh in particular industry.

Feeling of inability to control own production quickly (e.g., weather, in farming).

Belief in a "fair price," or a "fair profit."

Belief in "fair share" of market, or right to a "fair" living. (Note anti-chainstore legislation, control of number and kind of lawyers, doctors, barbers, etc.)

Community support of local concerns.

Understanding of exchange economy, especially in field of international trade.

Feeling of ability to set own prices, vs. compulsion to follow "price-leader" or "the market."

Belief in "free competition" as a basic and desirable element in the economy and in own industry.

Loyalty to industry association or cooperative.

Climate of business opinion (business tendency to seek controls in bad times and throw them off in good.)

Fear of court interpretation of "intention" in merging.

Belief in need for and desirability of governmental controls in own and other industries.

Belief in strength of, and attitude towards, economic behavior of other economic groups.

References: 2, 11, 12, 16, 30, 35, 38, 50, 68, 83, 167, 189, 217, 227, 259, 293.

H-1. *Area of decision to be taken by entrepreneur: what customers to sell to.*

Directly to users or consumers; to retailers (a few large or many smaller); through own retailers; to wholesalers; to jobbers; or through food stores, variety stores, drug stores, etc.

H-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Knowledge of current distribution practices and results.

Desire to promote own brand.

Fear of being at mercy of one or a few large buyers.

Belief that "fair" price will be paid.

Preference for or avoidance of sharp bargaining.

Need to retain goodwill of existing customers.

Loyalty to friends, neighbors, and other habitual customers.

Estimate of credit-worthiness of customers.

Prestige gained or lost by sale to particular users, retailers, etc.

Attitudes of compliance with governmental regulations (rationing, priorities, set–asides).

Attitude towards price-maintenance by dealers.

Attitude towards engaging in price competition.

I-1. *Area of decision to be taken by entrepreneur: choice of sources of supply and of capital.*

From what particular source of supply to buy raw materials, supplies, services, equipment, fuel, goods to resell, etc.; where to raise money for additional capital.

I-2. *Psychological factors likely to influence decision* (excluding economic data and expectations).

Habitual or traditional channels of supply.

Community pressures or individual attitudes for or against (a) particular social, political, religious, or racial groups, or (b) types of economic organization.

Reciprocity.
Loyalty to organization (cooperative, grange, etc.).
Friendship.
Desire to have dependable source of supply.
Dislike of dependence on too few sources.
Desire to control source of supply.
Fear of control by source of supply or by creditor.

J-1. *Area of decision to be taken by entrepreneur: whether and how to change wage levels or composition of labor force.*

Altering wage rates (without changing kind of rate); changing from hourly rate to piece rate, or vice versa; establishment or modification of bonus or incentive systems; changing wage structure (as by job analysis and classification); changing other components of compensation (vacations, pensions, etc.); hiring more employees or reducing labor force; changing proportion of particular classes of labor (e.g., special skills or unskilled labor replaced by new machinery); or varying hours of labor (over-time or part-time work, or change in regulation work-day) instead of numbers of employees.

J-2. *Psychological factors likely to influence decisions* (excluding economic data and expectations).

Community estimate of a "fair" wage, and community pressures to comply therewith (either by raising wages or by resisting labor's demands).

Industry estimates of a "fair" wage and industry pressures to comply therewith.

Personal relationships between owner or manager and labor force (as on farm, in small business establishment, family assistants in retail store, etc.).

Belief in "right" to freedom of managerial decision in wage questions.

Feeling of freedom of managerial decision (limited by governmental controls, labor strength and attitudes, community and industry pressures, etc.).

Traditional wage levels or differentials (e.g., low pay in banking).

Appraisal of importance of company to the community (services, payroll, etc.)

Feeling of responsibility for social impact of wage level or labor force decisions.

Understanding of needs and motives of employees, of unions, and of union leaders.

Belief in possibility of cooperative modification of wage rates or wage structure.

Knowledge of attitudes of labor force towards particular changes.

Belief in importance of keeping a going organization (especially key men) through bad times.

Expected reaction of employees to wage action or inaction.

References: **59, 60, 63, 76, 101, 122, 144, 145, 147, 213** (esp. pp. 285–287), **213a, 219.**

APPENDIX A: *Stratifying the Economic Sample*

In planning an empirical investigation of economic behavior, it is important to identify and hold constant as many as possible of the independent variables that are known or presumed to be able to influence the behavior studied. On investigation, it may turn out that some of them do not in fact influence this behavior, for the entrepreneur's knowledge of the situation is never complete, and much of what he knows he may ignore (**26**, **81**, **153**). Until it is known what elements really are important in given situations, empirical generalizations must be built up separately for highly homogenous and fully described strata (**14**, **26**, **101**).

The most important independent variables of a *general economic* nature that must be held constant or otherwise "partialed out" are: the levels of prices, the levels of wages, the volume of sales, the interest rate, the availability of particular commodities or services, the level (and kind) of taxes, the level of profits, business activity (in general, in the local community, in a given industry, and for a given company), the productivity of land, machinery, and labor, and the existence of war or other crises.

The most important *institutional* factors that must be taken into account in stratifying data are:

1. *Kinds of commodities or services produced or distributed.* The outstanding variables here, so far as decision-making is concerned are: (a) durable deferrable-purchase commodities (like automobiles) *vs.* non-durable (especially perishable) non-deferable-purchase commodities (such as cigarettes); (b) differentiated (usually brand-named) commodities (like soap) *vs.* standard commodities (like wheat); (c) consumers goods (like clothing) *vs.* producers goods (like industrial belting); (d) goods representing large purchases in a family or business budget (such as a machine) *vs.* goods whose purchase has little effect on a budget (like stationery) and which, in the case of consumers, is subject to "impulse" buying; (e) luxury commodities or services, especially if bought for prestige purposes, *vs.* essentials; (f) relative proportion of labor cost to total cost; (g) degree of individual craftsmanship.

2. *Size of company*, especially the number of people employed therein. Bureaucracy flourishes in industry as well as in government, and faces many of the same problems of diffusion of the decision-making process, red tape, passing the buck in risky decisions, inadequate communication between different elements, internal rivalries, jurisdictional disputes, etc.

3. *Spatial dispersion of company units*, and degree of responsibility allowed each unit. This factor may intensify the problems of a large company mentioned in (2) or may result in separate decision-making in many small, relatively independent units. The location of some units or operations outside the national boundaries of the country where the home office is located introduces, of course, new factors into the decision-making of local units.

4. *Stage of development of the industry.* A company in a "growth industry" (like television today) gives weight to quite different considerations from those considered important by a mature or declining industry (like black gunpowder manufacturing).

5. *Length of cycle of production.* In much farming several months, and in some plantation production (rubber, cocoa, bananas) several years, are needed before a decision to increase or reduce the scale of production can be made effective without disproportionate losses. Likewise, manufacture of locomotives or ships may involve a long cycle of production that markedly influences the psychology of decision-makers in that industry.

6. *Kind of market served.* Selling to many potential customers involves quite different considerations and tactics from selling to only a few. Where a substantial part of production is destined for export, additional considerations enter in. The practice of selling on organized exchanges (e.g., commodity exchanges, stock exchanges) or at auction also introduces special factors (**13**, **14**, **35**, **42**, **259**).

7. *Number of competitors, and domination of market by a company or by a competitor.*

This is of basic importance in appraising the likely effect of a price change on sales (**13, 14, 35, 42, 259, 271**).

8. *Ease or difficulty of entry of new competitors into same line of production or distribution.* In some industries, the great size of the initial investment needed, the existence of legal barriers, or the unavailability of patented processes may limit the possibilities of additional competition (**13, 14, 35, 42**).

9. *Extent of coordination of action as between companies.* Cooperation, parallel action, or collusion of sellers or buyers may characterize some market situations (**11, 12, 35, 38, 68, 83, 101, 217, 259, 271**).

10. *Separation of ownership and control.* The wide ownership of some companies or the inactivity of major ownership interests may result in most significant decisions being made by hired managers. The motives and other factors affecting their decisions may be quite different from those of owners (**23, 26, 62, 127a**).

11. *Extent of public regulation (or ownership and operation).* Where prices and/or wages are publicly regulated, where profits are limited or non-existent, where accounts are made public, managerial decisions may be quite different from those of "free" enterprises.

12. *Existence of economic coercion* (through patent ownership, debt, monopoly, etc.).

13. *Social controls other than legislation* (mores, local social pressures, etc.).

BIBLIOGRAPHY

1. Abramowitz, M. Monopolistic selling in a changing economy. *Quart. J. Econ.*, 1938, **52**, 191–214.
2. Allport, G., Murphy, G., & May, M. A. *Memorandum on research in competition and cooperation.* New York: Social Science Research Council, 1937.
3. Anderson, A. T., & Ross, R. C. Postwar farm jobs and farmers' purchase intentions. *Univ. Ill. Coll. Agri. Circ.* (Oct.) 1945, No. 592.
4. Anderson, B. M., Jr. *Social value.* New York: Houghton-Mifflin, 1911.
5. Anderson, B. M., Jr. *The value of money.* New York: R. R. Smith, 1936.
6. Anderson, Karl L. The unity of Veblen's theoretical system. *Quart. J. Econ.*, 1933, **47**, 598–626.
7. Andrews, P. W. S. A further enquiry into the effects of rates of interest. *Oxf. Econ. Pap.*, 1940, No. 3, 33–73.
8. Andrews, P. W. S. *Manufacturing business.* London: Macmillan, 1949.
9. Andrews, P. W. S. A reconsideration of the theory of the individual business. *Oxf. Econ. Pap.*, 1949, **1**, N.S., 54–89.
10. Angell, J. W. The components of circular velocity of money. *Quart. J. Econ.*, 1937, **51**, 224–272.
11. Arnold, T. W. *The folklore of capitalism.* New Haven: Yale Univ. Press, 1937.
12. Arnold, T. W. *Bottlenecks of business.* New York: Reynal & Hitchcock, 1940.
13. Bain, J. S. Market classifications in modern price theory. *Quart. J. Econ.*, 1942, **56**, 560–574.
14. Bain, J. S. Price and production policies. Ch. 4 in H. S. Ellis (Ed.), *A survey of contemporary economics.* Philadelphia: Blakiston, 1948.

14a. Bain, J. S. A note on pricing in monopoly and oligopoly. *Amer. econ. Rev.*, 1949, **39**, 448–464.

15. Baker, J. C. Executive compensation payments by large and small industrial companies. *Quart. J. Econ.*, 1939, **53**, 404–434.
16. Ballantine, O. O. Psychological bases for tax liability. *Harv. Bus. Rev.*, 1949, **27**, 200–208.
17. Barnard, C. I. *Functions of the*

executive. Cambridge: Harvard Univ. Press, 1939.

17a. Barnard, C. I. Functions and pathology of status systems. In W. F. Whyte (Ed.), *Industry and society.* New York: McGraw-Hill, 1946, 46–83.

18. Barnes, L. How sound were private postwar forcasts? *J. polit. Econ.*, 1948, **56**, 161–165.

19. Barnett, G. E. *Chapters on machinery and labor.* Cambridge: Harvard Univ. Press, 1926.

20. Bauer, P. T. Notes on cost. *Economica*, 1945, **12**, N. S., 90–100.

21. Becker, A. P. Psychological production and conservation. *Quart. J. Econ.*, 1949, **63**, 577–583.

22. Bentham, J. *Table of the springs of human action. The works of Jeremy Bentham*, Vol. 1. Edinburgh: William Tait, 1843. Pp. 195–219.

22a. Bentham, J. *A fragment on government and an introduction to the principles of morals and legislation.* New York: Macmillan, 1948. Chs. V and X.

23. Berle, A. A., & Means, G. C. *The modern corporation and private property.* New York: Macmillan, 1933.

24. Bloom, G. F. A note on Hicks' theory of invention. *Amer. econ. Rev.*, 1946, **36**, 83–96.

25. Blum, M. L. *Industrial psychology and its social foundations.* New York: Harper, 1949.

26. Boulding, K. E. The theory of the firm in the last ten years. *Amer. econ. Rev.*, 1942, **32**, 791–802.

27. Boulding, K. E. The incidence of a profits tax. *Amer. econ. Rev.*, 1944, **34**, 567–572.

28. Boulding, K. E. *The economics of peace.* New York: Prentice-Hall, 1945.

29. Boulding, K. E. A note on the theory of the black market. *Canad. J. econ. polit. Sci.*, 1947, **13**, 115–118.

30. Bradley, W. L. Taxation of cooperatives. *Harv. Bus. Rev.*, 1947, **25**, 576–586.

31. Brewster, J. M., & Parsons, H. L. Can prices allocate resources in American agriculture? *J. Fm. Econ.* 1946, 28, 938–960.

32. Bronfenbrenner, M. Applications of the discontinuous oligopoly demand curve. *J. polit. Econ.*, 1940, **48**, 420–427.

33. Brown, E. C. Some evidence on business expectations. *Rev. Econ. & Statist.*, 1949, **31**, 236–238.

34. Burns, A. F. *The cumulation of economic knowledge.* New York: National Bureau of Economic Research (28th Annual Report), 1948.

35. Burns, A. R. *The decline of competition.* New York: McGraw-Hill, 1936.

36. Butters, J. K., & Lintner, J. V. *Effect of federal taxes on growing enterprises.* Cambridge: Harvard Univ. Press, 1945.

36a. Campbell, A., & Katona, G. A national survey of wartime savings. *Pub. Opin. Quart.*, 1946, **10**, 373–381.

36b. Cartwright, D. Surveys of the war finance program. In *Proceedings of the Conference on Consumers' Interests.* Philadelphia: Univ. of Pennsylvania Press, 1947.

36c. Cartwright, D. Some principles of mass persuasion. *Hum. Rel.*, 1949, **2**, 253–267.

36d. Cartwright, D. Psychological economics. In *Experiments in Social Progress.* New York: McGraw-Hill, 1950.

37. Carver, T. N. The behavioristic man. *Quart. J. Econ.*, 1918, **33**, 195–200.

38. Chamberlin, E. G. *The theory of monopolistic competition.* Cambridge: Harvard Univ. Press, 1938.

39. Chase, S. *The proper study of mankind.* New York: Harper, 1948.

40. Clark, J. M. Economics and modern psychology. *J. polit. Econ.*, 1918, **26**, 1–30, 136–166. Reprinted as

Ch. 4 of J. M. Clark, *A preface to social economics*. New York: Farrar & Rinehart, 1936.

41. Clark, J. M. The socializing of theoretical economics. In R. G. Tugwell (Ed.), *The trend of economics*. New York: A. A. Knopf, 1924, 71–102.
42. Clark, J. M. Toward a concept of workable competition. *Amer. econ. Rev.*, 1940, **30**, 241–256.
43. Clark, J. M. Realism and relevance in the theory of demand. *J. polit. Econ.*, 1946, **54**, 347–353.
44. Clark, Pearl F. *The challenge of American know-how*. New York: Harper, 1948.
45. Coase, R. H. Some notes on monopoly price. *Rev. Econ. Stud.*, 1937, **5**, 17–31.
46. Cochran, T. C. A plan for the study of business thinking. *Polit. Sci. Quart.* 1947, **62**, 82–90.
47. Cooley, C. H. *Social process*. New York: Scribner's, 1918.
48. Cooper, W. W. Revisions to the theory of the firm. *Amer. econ. Rev.*, 1949, **39**, 1204–1222.
49. Copeland, M. A. Professor Knight on psychology. *Quart. J. Econ.*, 1925, **40**, 134–151.
50. Cox, R. Non-price competition and the measurement of prices. *J. Marketing*, 1946, **10**, 370–383.
51. Davenport, H. J. Scope, method and psychology in economics. *J. Phil. Psychol. sci. Meth.*, 1917, **14**, 617–626.
52. Davies, A. K. Sociological elements in Veblen's economic theory. *J. polit. Econ.*, 1945, **53**, 132–149.
53. Denis, H. La theorie psychologique de la formation des prix devant la critique contemporaine. *Rev. d'econ. polit.*, 1949, **56**, 166–182.
54. Dibblee, G. B. *The psychological theory of value*. London: Constable, 1924.
55. Dickinson, Z. C. The relations of recent psychological developments to economic theory. *Quart. J. Econ.*, 1919, **33**, 377–421.
56. Dickinson, Z. C. *Economic motives*. Cambridge: Harvard Univ. Press, 1922.
57. Dickinson, Z. C. Quantitative methods in psychological economics. *Amer. econ. Rev.*, 1924, **14** (Suppl.), 117–126.
58. Dickinson, Z. C. Economics and psychology. In W. F. Ogburn & A. Goldenweiser, *The social sciences and their interrelations*. New York: Houghton-Mifflin, 1927. Pp. 160
59. Dickinson, Z. C. *Compensating industrial effort*. New York: Ronald, 1937.
60. Dickinson, Z. C. *Collective wage determination*. New York: Ronald, 1941.
61. Douglas, P. H. The reality of non-commercial incentives in economic life. In R. G. Tugwell, (Ed.), *The trend of economics*. New York: A. A. Knopf, 1924.
62. Drucker, P. F. *Concept of the corporation*. New York: John Day, 1946.
63. Dunlop, J. T. *Wage determination under trade unions*. New York: Macmillan, 1944.
64. Durbin, E. F. M. Methods of research—a plea for cooperation in the social sciences. *Econ. J.*, 1938, **48**, 183–195.
65. Easterbrook, W. T. The climate of enterprise. *Amer. econ. Rev.*, 1949, **39** (Suppl.), 322–335.
66. Eban, A. S. Some social and cultural problems of the Middle East. *Internat. Affairs*, 1947, **23**, 367–375.
67. Ebersole, J. F. The influence of interest rates upon entrepreneurial decisions in business—a case study. *Harv. Bus. Rev.*, 1938, **17**, 35–39.
68. Edwards, C. D. Can the anti-trust laws preserve competition? *Amer. econ. Rev.*, 1940, **30** (Suppl.), 164–179.

69. EITEMAN, W. J. *Price determination; business practice versus economic theory.* Ann Arbor: Univ. of Michigan, Bur. of Bus. Research. Report No. 16, 1949.

70. ELLIS, H. S. Some fundamentals in the theory of velocity. *Quart. J. Econ.*, 1938, **52**, 431–472.

71. EVANS, G. H., JR. The entrepreneur and economic theory: a historical and analytical approach. *Amer. econ. Rev.*, 1949, **39**, 336–348.

72. EZEKIEL, M. *Methods of correlation analysis* (2nd Ed.). New York: Wiley, 1941.

73. EZEKIEL, M. Statistical investigation of saving, consumption, and investment: II. *Amer. econ. Rev.*, 1942, **32**, 272–307.

74. FELLNER, W. Monetary policies and hoarding in periods of stagnation. *J. polit. Econ.*, 1943, **51**, 191–205.

75. FELLNER, W. *Monetary policies and full employment.* Berkeley: Univ. of Calif., 1946.

76. FELLNER, W. Prices and wages under bilateral monopoly. *Quart. J. Econ.*, 1947, **61**, 503–532.

77. FELLNER, W. Average-cost pricing and the theory of uncertainty. *J. polit. Econ.*, 1948, **56**, 249–252.

78. FELLNER, W. Employment theory and business cycles. Ch. 2 in Amer. Econ. Assn., *A survey of contemporary economics.* Philadelphia: Blakiston, 1948.

79. FISHER, A. G. B. *The clash of progress and security.* London: Macmillan. 1935.

80. FISHER, I. Health and war. *Am. Lab. Leg. Rev.*, 1918, **8**, 9–20.

81. FLORENCE, P. S. *Economics and human behavior.* London: Kegan Paul, 1927.

82. FRIEDMAN, M., & KUZNETS, S. *Income from independent professional practice.* New York: National Bureau of Economic Research, 1945.

83. GALBRAITH, J. K. Monopoly and the concentration of economic power. Ch. 3 of Amer. Econ. Assn., *A survey of contemporary economics,* H. S. Ellis (Ed.), Philadelphia: Blakiston, 1948.

84. GAMBS, J. S. *Beyond supply and demand.* New York: Columbia Univ. Press, 1946.

85. GILBERT, R. V., & PERLO, V. The investment-factor method of forecasting business activity. *Econometrica*, 1942, **10**, 311–316.

86. GOODFELLOW, D. M. *Principles of economic sociology.* London: G. Routledge, 1939.

87. GORDON, R. A. Ownership and compensation as incentives to corporation executives. *Quart. J. Econ.*, 1940, **54**, 455–473.

88. GORDON, R. A. *Business leadership in the large corporation.* Washington: Brookings Inst., 1945.

89. GORDON, R. A. Short-period price determination in theory and practice. *Amer. econ. Rev.*, 1948, **38**, 265–288.

89a. GRIFFIN, C. E. *Enterprise in a free society.* Chicago: Richard D. Irwin, 1949.

89b. GRUCHY, A. G. *Modern economic thought: the American contribution.* New York: Prentice-Hall, 1947.

90. HAHN, F. H. A note on profit and uncertainty. *Economica*, 1947, **14**, 211–225.

91. HALEY, B. F. Value and distribution, Ch. 1 of Amer. Econ. Assn., *A survey of contemporary economics.* Philadelphia: Blakiston, 1948.

92. HALL, R. L., & HITCH, C. J. Price theory and business behavior. *Oxf. Econ. Pap.*, 1939, No. 2, 12–45.

93. HAMILTON, W. H. The institutional approach to economic theory. *Amer. econ. Rev.*, 1918, **8** (Suppl.) 309–318 & discussion 318–324.

94. HAMILTON, W. H., *et al. Price and price policies.* New York: McGraw-Hill, 1938.

95. HANSEN, A. H. *Full recovery or stagnation?* New York: Norton, 1938.

96. HANSEN, A. H. *Fiscal policy and*

business cycles. New York: Norton, 1941.

97. Harrod, R. F. Scope and method of economics. *Econ. J.*, 1938, **48**, 383–412.

98. Harrod, R. F. Price and cost in entrepreneurs' policy. *Oxf. Econ. Pap.*, 1939, No. 2, 1–11.

99. Hart, A. G. Anticipations, uncertainty, and dynamic planning. *J. Bus.*, 1940, **13** (Special Suppl.).

100. Hart, A. G. Risk, uncertainty, and the unprofitability of compounding probabilities. Univ. of Chicago Dept. of Economics, *Studies in Mathematical Economics and Econometrics.* Chicago: Univ. of Chicago Press, 1942.

100a. Hart, A. G. Assets, liquidity and investment. *Amer. econ. Rev.*, 1949, **39** (Suppl.), 171–181.

101. Hayes, S. P., Jr. Psychology of industrial conciliation and arbitration procedures. Ch. 18 in G. W. Hartman &T. M. Newcomb (Eds.), *Industrial conflict, a psychological interpretation.* New York: Cordon, 1939.

102. Hayes, S. P., Jr. The business cycle: psychological approaches. *Polit. Sci. Quart.*, 1948, **63**, 82–98.

103. Henderson, H. D. The significance of the rate of interest. *Oxf. Econ. Pap.*, 1938, No. 1, 1–13.

104. Hermann, Friedrich B. W. V. *Staatswirtschaftliche Untersuchungen* (Rev. Ed.). Leipzig: A. Lorentz, 1924.

105. Herring, P. The social sciences in modern society. *Soc. Sci. Res. Coun. Items*, 1947, **1**, 2–6.

106. Hicks, J. R. *Value and Capital.* New York: Oxford Univ. Press, 1938.

107. Hicks, J. R., & Allen, R. G. D. A reconsideration of the theory of value. *Economica*, 1934, N. S. No. **1–2**, 52–76, 196–219.

108. Higgins, B. E. Elements of indeterminacy in the theory of nonperfect competition. *Amer. econ. Rev.*, 1939, **29**, 468–479.

109. Higgins, B. E. The economic man and economic science. *Canad. J. econ. polit. Sci.*, 1947, **13**, 587–599.

110. Hobson, J. A. *Work and wealth.* New York: Macmillan, 1914. (Esp. Ch. 4).

111. Hurwicz, L. Theory of the firm and of investment. *Econometrica*, 1946, **14**, 109–136.

112. Hutchison, T. W. *The significance and basic postulates of economic theory.* London: Macmillan, 1938.

113. Hutchison, T. W. Reply to Professor Knight. *J. polit. Econ.*, 1941, **49**, 732–750.

114. Jevons, W. S. *The theory of political economy.* (4th Ed.) London: Macmillan, 1931.

114a. Jones, G. M. Why do businesses fail? *Dun's Rev.*, 1949 (Dec.), **57**, 17ff.

115. Kaplan, A. D. H. *Small business: its place & problems.* New York: McGraw-Hill, 1948.

116. Katona, G. *War without inflation, the psychological approach to problems of war economy.* New York: Columbia Univ. Press., 1942.

117. Katona, G. *Price control and business.* Bloomington, Ind.: Principia Press, 1945.

118. Katona, G. Psychological analysis of business decisions and expectations. *Amer. econ. Rev.*, 1946, **36**, 44–62.

119. Katona, G. Contribution of psychological data to economic analysis. *J. Amer. statist. Assn.*, 1947, **42**, 449–459.

119a. Katona, G. Analysis of dissaving. *Amer. econ. Rev.*, 1949, **39**, 673–688.

119b. Katona, G. Effect of income changes on the rate of saving. *Rev. econ. Statist.*, 1949, **31**, 95–103.

119c. Katona, G. Financial surveys among consumers. *Hum. Rel.*, 1949, **2**, 3–11.

120. Kaufmann, F. On the postulates of economic theory. *Soc. Res.*, 1942, **9**, 379–395.

121. Keirstead, B. S., & Coore, D. H. Dynamic theory of rents. *Canad. J. econ. polit. Sci.*, 1946, **12**, 168–172 (esp. footnote 4, p. 170).

122. Kennedy, V. D. *Union policy and incentive wage methods*. New York: Columbia Univ. Press, 1945.

123. Keynes, J. M. *The general theory of employment, interest and money*. London: Macmillan, 1936.

124. Kisselgoff, A. Liquidity preference of large manufacturing corporations, 1921–1939. *Econometrica*, 1945, **13**, 334–344.

125. Kitson, H. D. Economic implications in the psychological doctrine of interest. *J. polit. Econ.*, 1920, **28**, 332–338.

126. Klein, L. R. Macroeconomics and the theory of rational behavior. *Econometrica*, 1946, **14**, 93–108.

127. Klein, L. R. *Studies in investment behavior*. New York: National Bureau of Economic Research, Nov. 1949 (Mimeogr.).

127a. Knauth, Oswald. *Managerial enterprise: its growth and methods of operation*. New York: Norton, 1948.

128. Knight, F. H. *Risk, uncertainty & profit*. New York: Houghton-Mifflin, 1921.

129. Knight, F. H. The limitations of scientific method in economics. In R. G. Tugwell (Ed.), *The trend of economics*. New York: Knopf, 1924. Pp. 229–267.

130. Knight, F. H. Economic psychology and the value problem. *Quart. J. Econ.*, 1925 (May), **39**, 372–409.

131. Knight, F. H. Fact and metaphysics in economic psychology. *Amer. econ. Rev.*, 1925, **15**, 247–266.

132. Knight, F. H. Interest. *Encyclopaedia of the Social Sciences*, 8, 131–143. New York: Macmillan.

133. Knight, F. H. Economic science in recent discussion. *Amer. econ. Rev.*, 1934, **24**, 225–238.

134. Knight, F. H. The quantity of capital and the rate of interest. *J. polit. Econ.*, 1936, **44**, 433–463, 612–642.

135. Knight, F. H. Note on Dr. Lange's interest theory. *Rev. econ. Stud.*, 1937, **4**, 223–230.

136. Knight, F. H. Professor Parsons on economic motivation. *Canad. J. econ. polit. Sci.*, 1940, **6**, 460–465.

137. Knight, F. H. "What is truth" in economics. *J. polit. Econ.*, 1940, **48**, 1–32.

138. Knight, F. H. Realism and relevance in the theory of demand. *J. polit. Econ.*, 1944, **52**, 289–318.

139. Knight, F. H. Diminishing returns from investment. *J. polit. Econ.*, 1944, **52**, 26–44.

139a. Lachmann, L. M. The role of expectations in economics as a social science. *Economica*, 1943, N.S. **10**, 12–23.

140. Lange, O. The place of interest in the theory of production. *Rev. econ. Stud.*, 1936, **3**, 159–192.

141. Lange, O. *Price flexibility and employment*. Bloomington, Ind.: Principia Press, 1944.

142. Lange, O. The scope and method of economics. *Rev. econ. Stud.*, 1945–46, **13**, 19–32.

143. Lerner, A. P. *The economics of control*. New York: Macmillan, 1944.

144. Lester, R. A. Shortcomings of marginal analysis for wage-employment problems. *Amer. econ. Rev.*, 1946, **36**, 62–82.

145. Lester, R. A. Marginalism, minimum wages, and labor markets. *Amer. econ. Rev.*, 1947, **37**, 135–48.

146. Lester, R. A. Absence of elasticity considerations in demand to the firm. *Southern econ. J.*, 1948, **14**, 285–289.

146a. Lester, R. A. Equilibrium of the firm. *Amer. econ. Rev.*, 1949, **39**, 478–484.

147. Lester, R. A., & Robie, E. A. *Wages under national and regional collective bargaining: experience in seven industries*. Princeton: Princeton Univ. Press, 1946.

148. Lewisohn, S. A. Psychology in economics. *Polit. Sci. Quart.*, 1938, **53**, 233–238.

148a. Likert, R. The sample interview survey. In *Current trends in psychology*. Pittsburgh: Univ. of Pittsburgh Press, 1947, 196–225.

148b. Likert, R., & Katona, G. Relationship between consumer expenditures and savings: the contribution of survey research. *Rev. econ. Statist.*, 1946, **28**, 197–199.

149. Loewe, Adolph. *Economics and sociology*. London: G. Allen & Unwin, 1935.

150. Loewe, Adolph. Economic theory and social structure. *Manchester School*, 1936, **7**, 18–37.

151. Loewe, Adolph. A reconsideration of the law of supply and demand. *Soc. Res.*, 1942, **9**, 431–457.

152. Lutz, F. A. The interest rate and investment in a dynamic economy. *Amer. econ. Rev.*, 1945, **35**, 811–830.

153. Macaulay, F. R. *Some theoretical problems suggested by the movements of interest rates, bond yields and stock prices in the United States since 1856*. New York: National Bureau of Economic Research, 1938.

154. McCamy, J. L. Analysis of the process of decision-making. *Public Administration Rev.*, 1947, **7**, 41–48.

155. Machlup, F. Competition, pliopoly and profit. *Economica*, 1942, **9**, 1–23, 153–173.

156. Machlup, F. Marginal analysis and empirical research. *Amer. econ. Rev.*, 1946, **36**, 519–554.

157. Machlup, F. Rejoinder of an antimarginalist. *Amer. econ. Rev.*, 1947, **37**, 148–154.

158. Maclaurin, W. R. Wages and profits in the paper industry, 1929–39. *Quart. J. Econ.*, 1944, **43**, 196–228.

159. McMurry, R. N. The problem of resistance to change in industry. *J. appl. Psychol.*, 1947, **31**, 589–593.

160. Malthus, Thomas R. *Essay on the principle of population* (7th Ed.). London: Reeves & Turner, 1872.

161. Marot, Helen, *The creative impulse in industry*. New York: Dutton, 1918.

162. Marquis, D. G. Psychology of social change. In H. M. Dorr (Ed.), *Social implications of modern science. The Annals of the Amer. Acad. of Polit. and Social Sci.*, 1947, **249**, 75–80.

163. Marschak, J. Lack of confidence. *Soc. Res.*, 1941, **8**, 41–62.

164. Marschak, J., & Makower, H. Assets, prices and monetary theory. *Economica*, 1938, N.S. **5**, 261–288.

165. Marshall, A. *Principles of economics* (8th Ed.). London: Macmillan, 1922.

166. Mason, E. S. Price and production policies of large-scale enterprise. *Amer. econ. Rev.* (Proceedings), 1939, **19**, 61–74.

167. May, M. A., & Doob, L. W. *Competition and cooperation*. Soc. Sci. Rev. Counc. Bull. 25. New York: Social Science Research Council, 1937.

167a. Mayer, K. Toward understanding economic behavior. *Amer. J. Econ. Sociol.*, 1949, **8**, 321–335.

168. Meade, J. E., & Andrews, P. W. S. Summary of replies to questions on effects of interest rates. *Oxf. Econ. Pap.*, 1938, No. 1, 14–31.

169. Merton, R. K. *The expert and research in applied social science.* Notes for discussion (Mimeogr). New York: Columbia University, Bureau of Applied Social Research, Nov., 1947.

170. Metzler, L. A. The nature and stability of inventory cycles. *Rev. econ. Statist.*, 1941, **23**, 113–129.

171. Mill, J. S. *Auguste Comte and positivism.* Philadelphia: Lippincott, 1866.

172. Mitchell, W. C. The rationality of economic activity. *J. polit. Econ.*, 1910, **18**, 97–113, 197–216.

173. Mitchell, W. C. The backward art

of spending money. *Amer. econ. Rev.*, 1912, **2**, 269–281.

174. Mitchell, W. C. Human behavior in economics: a survey of recent literature. *Quart. J. Econ.*, 1914, **29**, 1–47.

175. Mitchell, W. C. The role of money in economic theory. *Amer. econ. Rev.*, 1916, **6** (Suppl.), 140–161.

176. Mitchell, W. C. Bentham's felicific calculus. *Polit. Sci. Quart.*, 1918, **33**, 161–183.

177. Mitchell, W. C. *Business cycles, the problem and its setting.* New York: National Bureau of Economic Research, 1928.

178. Mitchell, W. C. Empirical research and the development of economic science. In *Economic Research and the Development of Economic Science and Public Policy*. New York: National Bureau of Economic Research, 1946, 3–20.

179. Mitchell, W. C. *What happens during business cycles.* New York: National Bureau of Economic Research, 1949 (Mimeogr.).

180. Modigliani, F. *Expectations and business fluctuations.* Progress Reports (Mimeogr.). Urbana: Univ. of Illinois, March & Sept. 1949.

181. Moore, W. E. Primitives and peasants in industry. *Soc. Res.*, 1948, **15**, 44–81.

182. Moore, W. E. Theoretical aspects of industrialization. *Soc. Res.* 1948, **15**, 277–303.

183. Moulton, H. G. Some comments on research method. In *Economic Research and the Development of Economic Science and Public Policy.* New York: National Bureau of Economic Research, 1946, 37–52.

184. Neal, A. C. Marginal cost and dynamic equilibrium of the firm. *J. polit. Econ.*, 1942, **50**, 45–64.

185. Myers, C. O. Approaches and problems in wage research. *Amer. econ. Rev.*, 1947, **37**, 367–374.

186. Norris, Ruby Turner. *The theory of consumer's demand.* New Haven: Yale Univ. Press, 1941.

187. Nourse, E. G. *Price making in a democracy.* Washington: Brookings Inst., 1944. (Esp. 98–105.)

188. Nourse, E. G. Economics in the public service. *Amer. econ. Rev.*, 1947, **37**, 21–30.

189. Nourse, E. G., & Drury, H. B. *Industrial price policies and economic progress.* Washington: Brookings Inst., 1938.

190. Noyes, C. R. *Economic man.* New York: Columbia Univ. Press, 1948.

191. Ogburn, W. F. The psychological basis for the economic interpretation of history. *Amer. econ. Rev.*, 1918, **8** (Suppl.), 291–305.

192. Oliver, H. M., Jr. Average cost and long-run elasticity of demand. *J. polit. Econ.*, 1947, **55**, 212–221.

193. Oliver, H. M., Jr. Marginal theory and business behavior, *Amer. econ. Rev.*, 1947, **37**, 375–383.

194. Pantaleoni, M. *Pure economics,* New York: Macmillan, 1898.

195. Pareto, V. *The mind and society.* New York: Harcourt, Brace, 1935.

196. Parker, C. H. Motives in economic life. *Amer. econ. Rev.*, 1918, **8** (Suppl.), 212–231.

197. Parker, C. H. *The casual laborer and other essays.* New York: Harcourt Brace, 1920

198. Parsons, T. Wants & activities in Marshall. *Quart. J. Econ.*, 1931, **46**, 101–140.

199. Parsons, T. Some reflections on "The nature & significance of economics." *Quart. J. Econ.*, 1934, **48**, 511–545.

200. Parsons, T. Sociological elements in economic thought. *Quart. J. Econ.*, 1935, **49**, 414–453, 646–667.

201. Parsons, T. *The structure of social action.* New York: McGraw-Hill, 1937

202. Parsons, T. The motivation of economic activities. *Canad. J. econ. polit. Sci.*, 1940, **6**, 187–202.

203. Parsons, T. Note on institutionalism and the role of institutions. In H. E. Barnes, Howard Becker and Frances B. Becker, *Contemporary social theory*. New York: Appleton-Century, 1940. Pp. 642–646.
204. Parsons, T. Reply to Professor Knight. *Canad. J. econ. polit. Sci.*, 1940, **6**, 466–472.
205. Phelps, O. W. Theory of business communications. *J. Bus.*, 1942, **15**, 343–360.
206. Phillips, C. F. Major areas for marketing research. *J. Marketing*, 1946, **11**, 21–26.
207. Pigou, A. C. *Industrial fluctuations*. London: Macmillan, 1927.
208. Pinney, H. The institutional man. *J. polit. Econ.*, 1940, **48**, 543–562.
209. Pinney, H. The structure of social action. *Ethics*, 1940, **50**, 164–192.
210. Reddaway, W. B. Irrationality in consumers demand. *Econ. J.*, 1936, **46**, 419–423.
211. Reder, M. W. A reconsideration of the marginal productivity theory. *J. polit. Econ.*, 1942.
212. Reynaud, P. L. *Economie politique et psychologie experimentale*. Paris: Librairie de droit et de jurisprudence, 1946.
212a. Reynaud, P. L. Psychologie moderne et economie politique. *Rev. d'econ. polit.*, 1946, **56**, 81–95.
212b. Reynaud, P. L. Psychologie et biologie modernes en economie politique. *Rev. d'econ. polit.*, 1949, **59**, 228–239.
213. Reynolds, L. G. Economics of labor. Ch. 7 in H. S. Ellis (Ed.), *A survey of contemporary economics* (Amer. Econ. Assn.). Philadelphia: Blakiston, 1948.
213a. Reynolds, L. G. Toward a short-run theory of wages. *Amer. econ. Rev.*, 1948, **38**, 289–308.
214. Ricardo, D. *Principles of political economy and taxation*. New York: Dutton (Everyman's Ed.), 1933.
215. Robbins, L. C. *An essay on the nature and significance of economic science*. London: Macmillan, 1932.
215a. Robbins, L. C. Remarks on the relationship between economics and psychology. *Manchester School*, 1934, **5**, 80–101.
216. Robinson, Joan. *Economics is a serious subject*. Cambridge: Heffer & Sons, 1932.
217. Robinson, Joan. *The economics of imperfect competition*. London: Macmillan, 1933.
218. Rose, J. R. Business education in a university. *J. Bus.*, 1947, **4**, 183–190.
219. Ross, A. M. The dynamics of wage determination under collective bargaining. *Amer. econ. Rev.*, 1947, **37**, 793–822.
220. Rothschild, K. W. Price theory and oligopoly. *Econ. J.*, 1947, **57**, 299–320.
221. Samuelson, P. A. A note on the pure theory of consumers' behaviour. *Economica*, 1938, **5**, 61–62.
222. Samuelson, P. A. Dynamics, statics, and the stationary state. *Rev. econ. Statist.*, 1943, **25**, 58–68.
222a. Samuelson, P. A. *Foundations of economic analysis*. Cambridge: Harvard Univ. Press, 1947.
223. Saxton, C. C. *Economics of price determination*. New York: Oxford Univ. Press, 1942.
224. Sayers, R. S. Business men and the terms of borrowing. *Oxf. Econ. Pap.*, 1940, No. 3, 23–31.
225. Schultz, H. *The theory and measurement of demand*. Chicago: Univ. of Chicago Press, 1938.
226. Schumpeter, J. A. *Business cycles*. New York: McGraw-Hill, 1939.
227. Schumpeter, J. A. *Capitalism, socialism and democracy* (2nd Ed.). New York: Harper, 1947.
228. Schumpeter, J. A. Vilfredo Pareto (1848–1923). *Quart. J. Econ.*, 1949, **63**, 147–173.
229. Scitovszky, T. A note on profit

maximization and its implications. *Rev. econ. Studies*, 1943, **11**, 57–60.

230. SELTZER, L. H. Is a rise in interest rates desirable or inevitable? *Amer. econ. Rev.*, 1945, **35**, 831–850.
231. SHACKLE, G. L. S. The expectational dynamics of the individual. *Economica*, 1943, N.S. **10**, 99–129.
232. SHACKLE, G. L. S. *Expectation in economics.* Cambridge, England: Cambridge Univ. Press, 1949.
233. SIMON, H. A. *Administrative behavior—a study of decision-making processes in administrative organization.* New York: Macmillan, 1947.
234. SLICHTER, S. H. Present savings and postwar markets. *Harv. Bus. Rev.* 1943, **22**, No. 1 Suppl.
235. SLICHTER, S. H. The state of economics. *Soc. Sci. Res. Coun. Items*, 1949, **3**, 25–28.
236. SMITH, A. *The theory of moral sentiments* (New Ed.). London: G. Bell, 1892.
237. SNOW, A. J. Psychology in economic theory. *J. polit. Econ.* 1924, **32**, 487–496.
238. SOUTER, R. W. The nature and significance of economic science in recent discussion. *Quart. J. Econ.*, 1933, **47**, 377–413.
239. SOUTER, R. W. *Prolegomena to relativity economics.* New York: Columbia Univ. Press, 1933.
240. SPIEGEL, H. W. The war economy and economic man. *J. Bus.*, 1943, **16**, 1–6.
241. STAEHLE, H. The measurement of statistical cost functions, an appraisal of some recent contributions. *Amer. econ. Rev.*, 1942, **32**, 321–333.
242. STERN, J. Resistances to the adoption of technological inventions. In United States National Resources Committee, Science Committee, *Technological trends and national policy.* Washington: U. S. Government Printing Office, 1937, 39–66.
243. STIGLER, G. J. Notes on the theory of duopoly. In Amer. Econ. Assn. *Readings in the theory of income distribution*, 530–531. Philadelphia: Blakiston, 1946.
244. STIGLER, G. J. The kinky oligopoly demand curve and rigid prices. *J. polit. Econ.*, 1947, **55**, 432–437.
245. STIGLER, G. J. Professor Lester and the marginalists. *Amer. econ. Rev.*, 1947, **37**, 154–157.
246. STUART, H. W. Hedonistic interpretation of subjective value. *J. polit. Econ.*, 1895, **4**, 64–84.
247. TAUSSIG, F. W. *Inventors and money-makers.* New York: Macmillan, 1930.
248. TAUSSIG, F. W., & JOSLYN, C. S. *American business leaders.* New York: Macmillan, 1932.
249. TAWNEY, R. H. *The acquisitive society.* New York: Harcourt, Brace, 1920.
250. TAYLOR, O. H. Economic theory and certain non-economic elements in social life. Ch. II of Part III, of *Explorations in economics*, 380–390. New York: McGraw-Hill, 1936.
251. TEAD, ORDWAY, *Instincts in industry, a study of working-class psychology.* New York: Houghton Mifflin, 1918.
252. TERBORGH, G. W. *The bogey of economic maturity.* Chicago: Machinery & Allied Products Inst., 1945.
253. THOMPSON, G. C. *Forecasting sales.* New York: National Industrial Conference Board, 1947 (Studies in Business Policy No. 25).
254. THURSTONE, L. L. The indifference function. *J. soc. Psychol.*, 1931, **2**, 139–167.
255. THURSTONE, L. L. The prediction of choice. *Psychometrika*, 1945, **10**, 237–253.
256. TINBERGEN, J. The notions of horizon and expectancy in dynamic economics. *Econometrica*, 1933, **1**, 247–264.
257. TINBERGEN, J. *The reformulation of*

current business cycle theories as refutable hypotheses. New York: National Bureau of Economic Research, Nov. 1949 (Mimeogr.).

258. TOBIN, J. Liquidity preference and monetary policy. *Rev. econ. Statist.*, 1947, **29**, 124–131.

259. TRIFFIN, R. A. *Monopolistic competition and general equilibrium theory.* Cambridge, Mass.: Harvard Univ. Press, 1940.

260. TUGWELL, R. G. Human nature in economic theory. *J. polit. Econ.*, 1922, **30**, 317–345.

261. TUGWELL, R. G. Experimental economics. In R. G. Tugwell (Ed.), *The trend of economics*, 369–422. New York: Knopf, 1924.

262. ULMER, M. J., & NIELSEN, ALICE. Business turnover and causes of failure. *Survey of Current Business*, 1947, **27**, No. 4, 10–16.

263. VEBLEN, T. V. Why is economics not an evolutionary science? *Quart. J. Econ.*, 1898, **12**, 373–397.

264. VEBLEN, T. V. *The theory of the leisure class: an economic study of the evolution of institutions.* New York: Macmillan, 1899.

265. VEBLEN, T. V. The preconceptions of economic science. *Quart. J. Econ.* 1899, **13**, 121–150; 396–426; & 1900, **14**, 240–269.

266. VEBLEN, T. V. *The theory of business enterprise.* New York: Scribner's, 1904.

267. VEBLEN, T. V. The limitations of marginal utility. *J. polit. Econ.*, 1909, **17**, 620–636.

268. VEBLEN, T. V. *The instinct of workmanship and the state of the industrial arts.* New York: Macmillan, 1913.

269. VEBLEN, T. V. *The place of science in modern civilization.* New York: Viking, 1919.

270. VILLARD, H. H. Monetary theory. Ch. 9 in H. S. Ellis (Ed.), *A survey of contemporary economics* (Amer. Econ. Assoc.). Philadelphia: Blakiston, 1948.

271. VON NEUMANN, J., & MORGENSTERN, O. *The theory of games and economic behavior.* Princeton: Princeton Univ. Press, 1944.

272. WALKER, E. R. *From economic theory to policy.* Chicago: Univ. of Chicago Press, 1943.

273. WALKER, K. F. The psychological assumptions of economics. *Econ. Record* (Melbourne), 1946, **22**, 66–82.

274. WALLICH, H. C. The changing significance of the interest rate. *Amer. econ. Rev.*, 1946, **36**, 761–787, esp. 765–769.

275. WALLIS, W. A., & FRIEDMAN, M. The empirical derivation of indifference functions. *Studies in mathematical economics and econometrics*, 175–189. Chicago: Univ. Chicago Press, 1942.

276. WARBURTON, C. The misplaced emphasis in contemporary business fluctuation theory. *J. Bus.*, 1946, **19**, 199–220.

277. WEBER, M. *The Protestant ethic and the spirit of capitalism.* (Trans. by T. Parsons.) London: G. Allen & Unwin, 1930. (Esp. Ch. II.)

278. WEISSKOPF, W. A. Psychological aspects of economic thought. *J. polit. Econ.*, 1949, **57**, 304–314.

278a. WEISSKOPF, W. A. Individualism and economic theory. *Amer. J. Econ. Sociol.*, 1950, **9**, 317–333.

279. WHITEHEAD, T. N. Social motives in economic activities. *Occupational Psychol* (London), 1938, **12**, 271–290.

280. WHITMAN, R. M. Demand functions for merchandise at retail. In O. Lange *et al.* (Eds.), *Studies in mathematical economics and econometrics* (Henry Schultz Memorial Volume); 208–221. Chicago: Univ. Chicago Press, 1942.

281. WICKSTEED, P. H. *The common sense of political economy, including a study of the human basis of economic law.* London: Macmillan, 1910.

282. Wieser, F. *Das Gesetz der Macht.* Vienna: J. Springer, 1926.

283. Wolfe, A. B. Functional economics. In R. G. Tugwell (Ed.), *The trend of economics*, 443–482. New York: Knopf, 1924.

284. Wolfe, A. B. Sociology and economics. In W. F. Ogburn and A. Goldenweiser, *The social sciences and their inter-relations*, 299–310. Boston: Houghton Mifflin, 1927.

285. Wolfe, A. B. Economy and democracy. *Amer. econ. Rev.*, 1944, **34**, 1–20.

286. Wooton, Barbara F. *Lament for economics.* London: Allen & Unwin, 1938.

287. Working, H. The investigation of economic expectations. *Amer. econ. Rev.*, 1949, **39** (Suppl.), 150–166, and following discussion.

288. Wright, D. McC. Professor Knight on limits to the use of capital. *Quart. J. Econ.* 1944, **58**, 331–358.

289. Wright, D. McC. *The economics of disturbance.* New York: Macmillan, 1947. (Esp. Ch. 4.)

290. Yntema, T. O. Competition as a norm of economic behavior. *J. Bus.*, 1941, **14**, 270–283.

291. *Executive forecast.* (Anon.) (Mail survey of business expectations.) *Fortune*, Jan. 1950 and semi-annually previously, beginning Feb. 1947.

292. *Plant and equipment outlays.* (Anon.) (Mail survey of business plans.) *Survey of Current Business*, 1949, **29**, No. 1 (Oct.), 5 ff. (and similar quarterly reports since 1945).

293. *Why the urge to merge.* (Anon.) *Bus. Wk.*, 1948 (Sept. 4), 25–28.

294. *Cost behavior and price policy.* National Bureau of Economic Research (conference on price research), Price Studies No. 4. New York: National Bureau of Economic Research, 1943.

295. *Experience with human factors in agricultural areas of the world.* U. S. Dept. Agr. Extension Service and Office of Foreign Agricultural Relations, 1949.

296. *An exploration of factors motivating hog farmers in their production and marketing.* U. S. Dept. Agr. Bureau of Agricultural Economics, Aug. 1947.

297. *Progress report* (Jan., 1948) and *Second annual report* (Jan., 1949). Ann Arbor: Univ. of Michigan Institute for Social Research, Survey Research Center.

298. *Readings in business cycle theory.* American Economic Association. Philadelphia: Blakiston, 1944. Bibliographies on (a) Monetary interest theories (471–474); (b) Psychological influences, expectations and investment decisions (475–476); and (c) Business cycle control (486–487).

299. *The social sciences: their relations in theory and teaching.* Institute of Sociology and the International Student Service, Kings College. London: LePlay House Press, 1936.

300. *Surveys of business expectations.* Dun & Bradstreet, Inc. Include:
    Survey on business expectations and government policies, prepared for the Joint Committee on the Economic Report, May 23, 1947.
    Current business opinion, *Dun's Rev.*, June, 1948, 25–26.
    What business men expect in the last half of 1949, *Dun's Rev.*, June 1949, 11–12, and monthly surveys since then.

301. *1949 survey of consumer finances.* Board of Governors of the Federal Reserve System. Washington: 1949. (Reprinted from *Federal Reserve Bull.*, June–Oct., 1949.)

Received February 14, 1950.

THE MISUSE OF CHI-SQUARE—A REPLY TO LEWIS AND BURKE

CHARLES C. PETERS
Pennsylvania State College

In the November, 1949, issue of the *Psychological Bulletin*,[1] Lewis and Burke discuss "The Use and Misuse of the Chi-square Test" in a long article in which they allege misuses in all but three out of 14 articles, and charge the Peters and Van Voorhis textbook with six misses out of seven tries. In our case every one of these six criticisms is invalid, although they could have made, but did not, a seventh criticism which would have been valid. If the same distribution of validities holds for the other writers (a possibility that I have not investigated), the situation may not be so bad as they make out. Lewis and Burke play up only some partial applications of chi-square at the expense of more fundamental and general considerations, thus frequently overshooting the mark in their generalizations. They make no mention of Karl Pearson's 1900 article in the London-Edinburgh-Dublin *Philosophical Magazine* (Vol. **50**, pages 157–175), in which chi-square was first announced; yet no one can expect to have a basic understanding of this statistic without a thorough reading of this article and of Pearson's follow-up articles in volumes 14, 24 and 26 of *Biometrika*. We shall take up in turn some wrong allegations in the Lewis and Burke article.

1. *That chi-square applies fundamentally only to frequencies.* It is more general than that. The starting point for the derivation of the distribution function of chi and chi-square, which is always assumed for the "makings" of these statistics, is the normal distribution function for standard scores:

$$df = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz \qquad [i]$$

By methods recently available, the distribution function for chi and for chi-square come rather easily from this normal distribution function. They are:

$$df(\chi) = \frac{1}{\frac{n-2}{2}!\,2^{(n-2)/2}} \chi^{n-1} e^{-\chi^2/2} d(\chi) \qquad [ii]$$

---

[1] Lewis, Don, & Burke, C. J. The use and misuse of the chi-square test. *Psychol. Bull.*, 1949, **46**, 433–489.

$$df(\chi^2) = \frac{1}{\frac{n-2}{2}\,!\,2^{n/2}}\,(\chi^2)^{(n-2)/2}e^{-x^2/2}d(\chi^2) \qquad [iii]$$

where n is the number of degrees of freedom.[2]

In this derivation no conditions whatever are imposed about the nature of z except that it shall be a deviation from a true value, that the z's shall be normally distributed about this central value in the population, and that the deviations of which they are composed be divided by the true variance. The deviations may be individual variates or frequencies or means or any other statistics, provided only that they are normally distributed in the population, are from a true central value, and are divided by the true variance. However, aside from the case of frequencies, we can not often know the true variance and (somewhat less rarely) the true central value; but sometimes we can.

2. *That the sum of the theoretical frequencies must always equal the sum of the observed frequencies, and that our example reproduced in Table 1 of their article violates this principle.* The generalized definition of chi-square is:

$$\chi^2 = \sum \frac{(\mu - x)^2}{\tilde{\sigma}^2} \qquad [iv]$$

where μ is the true central value and the tilde over the σ means that it is the true standard deviation. The true variance for the frequency in a cell is $N\tilde{p}\tilde{q}$, where $\tilde{p}$ is the true proportion belonging in the cell and $\tilde{q}$ equals $(1-\tilde{p})$. So the generalized formula in the case of frequencies would be:

$$\chi^2 = \sum \frac{(f_t - f_o)^2}{N\widetilde{pq}}. \qquad [v]$$

This involves nothing whatever about what the frequencies shall sum to. In a setup such as ours of Table 1 in the Lewis and Burke article the f_o could be made to sum to equality with the f_t only by artificial doctoring, except in very rare cases. Yet, so long as the "cramping" features about to be discussed are not present, $[v]$ is the basic formula to apply to get the correct chi-square.

But Pearson showed, in the original article referred to above, that *if* the condition is imposed upon a universe of samples that the total N shall always be the same, and the same for theoretical and observed frequencies so that the sum of the errors is zero, then a system of intercorrelations is introduced of which account must be taken in getting the values for chi-square. For under these restrictions, an excess of frequencies in one cell must necessarily be accompanied by a deficiency in others. He set up the system of intercorrelations for this condition along with the squared deviations, worked them through an elaborate and highly technical solution by determinants, and emerged with the now well-known formula:

---

[2] Lewis and Burke do not give quite correctly the distribution function for chi-square. In their terminology, the exponent of chi-square should be $(r-2)/2$ instead of $(r+2)/2$.

$$\chi^2 = \sum \frac{(f_t - f_o)^2}{\tilde{p}N} = \sum \frac{(f_t - f_o)^2}{f_t}. \qquad [vi]$$

This is the formula that *must* be used when the limitation is imposed that the frequencies in a set shall always sum to N so that intercorrelations are involved; but, just as unequivocally, it must *not* be used otherwise.[3]

3. *That the frequencies of "non-occurrence" must always be present.* There is by no means always such necessity; it depends upon the nature of one's problem. One of the studies to which violation of this is charged is that by Grant and Norris, in which these investigators fitted a logarithmic curve and tested the agreement of the theoretical frequencies distributed to categories under it with the same aggregate of frequencies distributed according to observation. They applied the chi-square test in the customary manner for such frequency distribution testing. In such a procedure it would be entirely meaningless to set up for each category the number *not* in the category and compute chi-square elements for this second set of deviations to add to the ones conventionally used.

In their footnote on p. 435, Lewis and Burke also charge us with violation of this principle. In our example, statistics were given for the number of Italians, Russians, Poles, and Others who had been naturalized out of totals for those nationals in a community. The p was inferred from the sample, and the distribution tested for possible nationality differences. The chi-square was computed from only the row of "successes" by formula [*vi*]. We made the customary assumption that the total for the row is to be constant from sample to sample, and our generalization was about a universe of samples of which that description would hold. For our assumption of a fixed marginal total for the row, our procedure was in exact conformity with the conditions for which Pearson developed formula [*vi*]: a fixed N (the total number of naturalized) in a series of samples, f_t equal to f_o so that the sum of the errors in each sample is zero, a shifting incidence of the fixed number of naturalizations to the different nationalities in successive samples with the inevitable intercorrelations that would result, deviations of the frequencies in each cell from a theoretical number, and a test whether the deviations in this sample exceed those which could fall within the credible limits of sampling fluctuation. The degrees of freedom for this universe of samples are 3 and the p is .0024.

---

[3] We made a slip in using the customary formula, [*vi*], in connection with our dice throwing example when we should have used formula [*v*]. If formula [*v*] is used, chi-square is 12 with the same number of degrees of freedom. If Lewis and Burke had pointed this out, it would have been a valid criticism—the only one I acknowledge as legitimate.

To install in this table a row of "non-occurrences" (the non-naturalized) would be essentially to establish a series of ratios, one for each nationality group, apply formula [*vi*] to each with N being the number in that nationality group to get a chi-square for it, then sum across these two-cell columns for a composite chi-square for testing the homogeneity of these ratios. Karl Pearson says (*Biometrika*, Vol. **14**, p. 418, and Vol. **24**, p. 353) that a table thus set up is not a true contingency table to which the chi-square test is appropriate. At any rate, it automatically freezes the N_c's for the universe of samples, which is sometimes a natural and therefore desirable feature and sometimes is not. Instead of following blindly a rule regarding the nature of the universe about which one shall generalize, the researcher should make a choice of such universe according to what is most meaningful in his particular problem, then say what his universe is. In our naturalization problem, for example, we can state what the probability is that as great discrepancies in ratio as we found in our sample would arise merely by sampling fluctuation in a universe of samples of size N taken at random from the same or similar total communities; that $P(df=7)$ would be .000138. Or we could state what would be the probability of the obtained discrepancies in a sub-universe of samples with always exactly the same numbers of Poles, Russians, etc. and exactly the same aggregate number of naturalizations and non-naturalizations; that $P(df=3)$ is .000002. There is a whale of a difference, although here both, of course, are statistically highly significant refutations of the null hypothesis. In cases like this one the former generalization would seem to be much the more natural and meaningful. It requires, of course, that the values obtained from the sample be taken as acceptable estimates of the true values; but Pearson argued, very persistently and to this writer very convincingly, that these estimates are quite satisfactory and that the gain from using them is greater than the loss; and he supports his arguments with experimental evidence. An awkward and stilted generalization is too great a price to pay for excessive finicalness about true values. However, the fourfold-table (if used) is an exception, if one wishes outcomes that are comparable with those of differences between proportions or of tetrachoric correlations; for the standard error formulas of both of these statistics assume fixed marginal totals.

4. *That chi-squares may not be summed where the same set of individuals is used in the series, because the sets are not independent; and that this principle is violated in our dice throwing example of their Table 1.* That idea rests upon a misunderstanding of the nature of independence.

When you throw a set of dice, they fall according to some law actually *within* them, within the limits of sampling fluctuation involved in that law. You also know the credible limits in the distribution of samples according to some *external* law which you suspect may be the law operating within the sample. You apply a chi-square test to see whether the behavior of the sample lies within the credible limits according to the external law. If it does not, you reject the hypothesis that your external law may be the law operating in the sample. You have made this test for one cell. Then you make a second throw and in it the dice behave again according to the law within them, quite independently of their behavior during the first throw. And so on with successive throws. The reason for summing these is merely to increase the reliability of the test and thus get a more dependable generalization for *these* dice. To do anything else than use the same dice, or the same subjects, would rob your investigation of all meaning for testing that hypothesis. But if on a second throw you picked up only *some* of the dice while allowing the others to lie and be counted in the next throw as they lay in the former throw, the sets would not then be independent. So if you are intending to generalize about the bias of *these* dice, or *these* subjects in a guess-heads or a choice-of-values experiment, it is correct to sum successive throws of the same dice, or successive responses of the same subjects. But if your purpose is to generalize about a wider universe of dice or of subjects, then a new sample of that wider universe must be chosen on each successive trial. Then you can generalize about the wider universe from which the samples have been randomly drawn. Our generalization was specifically about *these* dice.

With dice this procedure could scarcely go awry. But, with human or animal subjects, if something happened between trials that might affect the next ones (as information about success), the chi-squares could not be meaningfully summed because we would have on the different trials not further samples from the same universe but each time a sample from a different universe. To study thus in a series of summed trials the bias of just *these* dice or *these* students may not be a very useful form of research, because the universe is too narrow to be, ordinarily, of much social importance. But it is sound statistically if one's purpose requires it.[4]

5. *That the view that the number of degrees of freedom appropriate in testing the fit of a normal curve depends upon the hypothesis to be tested is erroneously attributed to Karl Pearson.* It is not erroneously attributed

---

[4] The example in our dice throwing of Table 1 was not, of course, the report of a research. It was intended merely as a schematic example for illustrating the meaning of deviations from a true value, worked up in the conventional manner, excessively simplified for pedagogical purposes. The objection that it would be a poor research, and that the frequencies are too small, seems under the circumstances to be picayune.

to Pearson; it is correctly attributed to him, as a reading of *Biometrika*, Vol. **24**, pp. 359 and 361 will show.

6. *That certain writers, including us, recommend a chi-square test of goodness of fit of regression lines which is a rather poor one.* This should have been attributed to R. A. Fisher, not to either of the two writers to whom reference is made. We merely passed on Fisher's scheme, though unfortunately with our blessing. But in a footnote on the same page we referred the reader to "a better test of the goodness of fit of regression lines" in a section of our book where a test is developed in terms of the unbiassed correlation ratio, which is a statistic closely related to F.

Much of the difficulty with chi-square arises from a frenzy to extend its use into areas where it is not needed, because we have for them better statistics. All applications of chi-square with one degree of freedom should be cleared out with one sweep of the broom. For with one degree of freedom the distribution of chi is normal, so that nothing remains at that level that is distinctive for chi-square. Lewis and Burke say that for one *df* the square-root of chi-square (which is chi) is distributed as Student's t with N equal infinity. That is true because the distribution of Student's t with N equal to infinity *is* normal; but that fact does not follow from the stated distribution function of Student's t, because that contains no χ. But it does follow easily from our formula [*ii*]; by substituting 1 for n and remembering that χ^0 equals 1 and that $(-\frac{1}{2})!$ (which equals gamma $\frac{1}{2}$) equals $\sqrt{\pi}$, formula [*ii*] reduces to formula [*i*] (with, however, 2 in the numerator instead of 1, which does not affect the *shape* of the distribution and which compensates for the fact that chi's run only from zero to plus infinity instead of from minus infinity to plus infinity). If one applies in a four-fold table the well-known technique of a difference of proportions divided by the conventional standard error of a difference of proportions and goes to the normal curve for the interpretation of the probability, he will get the same result as by the chi-square technique to the hundredth decimal place—provided he is testing the same hypothesis and remembers that there belong in the standard error formula the "true" values of p and q. Even amateur statisticians know that you may not carry ratios from proportion to the normal curve and claim exactness for your probability in researches with such N's as we ever get in practice. But these same persons will go blithely and confidently ahead with chi-square, not knowing that none of the chi-square table entries hold strictly for N's short of infinity, and that this principle is violated just as badly in chi-square as it is in proportion.[5]

---

[5] There is much talk about what minimum number of frequencies is required in a cell before using it in a chi-square calculation: whether five or ten or twenty, or what. The number required for *exactness* in fulfilling the assumptions is infinite; for *any* number

Tetrachoric correlation, which can be applied to any four-fold table where the variates make a continuous distribution, also has a constructive meaning which shows the nature and the strength of the law that is present (while chi-square shows only whether *some* law is present without indicating either its nature or its strength). Tetrachoric r is, by the cosine-pi formula, as easily computed as chi-square; it assumes a normal distribution only over the table as a whole, not in the separate cells; and it has a good and reasonably simple test of reliability.

There are also available now more convenient tests than chi-square for testing the goodness of fit of the normal curve (because simpler), and for testing the goodness of fit of all regression lines. There remains little more for chi-square than the test of multiple contingency tables with greater complexity than four-fold. If research workers would not, in a mad rush to use chi-square, raise an unnecessary dust, they would not need to complain that they can not see.

This note has involved criticism of six allegations made by Lewis and Burke in their article. But, in spite of these particular criticisms and possibly some others that might have been made, their article contains many sound and useful ideas, and serves a good purpose. But it should be read with critical alertness, not taken offhand as gospel.

Received December 27, 1949.

---

short of infinity the determination is only rough, but good enough. In the binomial $(p+q)^N$, where $p+q$ equals 1, it is only as N (the frequency in the sample) approaches infinity that the distribution approaches the normality assumed for chi-square. When p (the proportion of the total N of the whole sample that belongs in the cell) and q are unequal, the distribution remains markedly skew far beyond the number usually mentioned as the minimum; but for any *definite* p it always approaches normality (around p) as the N of its exponent approaches infinity (and hence as the f in each cell approaches infinity). The damaging fact is not so much the intrinsic smallness of the f_t of the cell as it is the smallness of the p representing the proportion of the total number in the sample that belongs in the cell.

SOME COMMENTS ON "THE USE AND MISUSE OF THE CHI-SQUARE TEST"

NICHOLAS PASTORE
Hunter College

Lewis and Burke in their recent article in this journal[1] rightly call attention to the necessity for familiarity and understanding of the assumptions underlying the chi-square test. It is unfortunate, however, that some of the crucial points in this rather lengthy article are either incorrect or confusing. The purpose of this comment is to indicate some of these incorrect or confusing statements.

1. There is a typographical error in the statement of the equation of the distribution function of chi-square (p. 439, equation 8). The numerator of the exponent of χ^2 should be $(r-2)$ instead of $(r+2)$.

2. On p. 434f. Lewis and Burke discuss a number of errors which were made by Peters and Van Voorhis in illustrating a chi-square problem. In this discussion Lewis and Burke fail to present or indicate a possibly superior method for dealing with the problem. Such a method would involve the determination of the proportion of the fourteen throws which yield exactly 0, 1, 2, $\cdots$,12 aces. The proportions are the successive terms of the expansion of $14(q+p)^{12}$, where $p=1/6$ and $q=5/6$. The number of degrees of freedom would then be 11. (In the Peters and Van Voorhis example, however, the chi-square test is not applicable, as Lewis and Burke themselves point out, because the theoretical frequencies are too small.)

3. In connection with the previous example, Lewis and Burke raise another objection to the procedure followed by Peters and Van Voorhis: ". . . the observed frequencies lack independence. They lack independence because the same twelve dice were thrown each time" (p. 434). This is an irrelevant objection. The usual assumption is that the wear and tear on a set of twelve dice are unrelated to the probability of the specified event. Furthermore, if the purpose is to determine whether a particular set of dice is biased, the thing to do is to experiment with the given set of dice. Lewis and Burke imply that the set of dice should be changed with each successive throw.

4. On pp. 441–443 the authors present and discuss a presumably "incorrect" application of the additive property of chi-square. Actually the application of the additive property of chi-square to this example is correct.

Ninety-six students were each asked to guess the fall of a coin five times. The actual turn of the coin was never revealed. "The hypothesis under test is that the guesses of students *on five successive tosses* of a coin are purely chance occurrences, with the probability of a guess of heads (by any student on any toss) equal to the probability of a guess of tails" (p. 442).[2] In other words, the authors propose to study the response-tendencies of students in order to determine whether such response-tendencies represent the same type of occurrence as tossing a set of five unbiased coins 96 times. It is assumed that each set of coins is properly shuffled before each toss. On the basis of the hypothesis, and the assumed independence of each toss, it is possible to calculate the expected number of heads for each set of 96 tosses. Since the coin is assumed to be unbiased the expected number of heads is 48. On the basis of this theoretical picture it is possible to calculate the probability of obtaining either a set of observed frequencies, or a set of more divergent frequencies. This probability can be calculated for each of the five coins. The authors calculate the chi-square value for each coin and then determine the appropriate probability with one degree of freedom. (See Table 3, p. 442). The relevant question is whether the five values of chi-square can be added, as well as the degrees of freedom, in order to determine the over-all probability for the set of five chi-squares. The authors claim that this addition is incorrect because "between tosses, the five guesses of each student were undoubtedly interrelated" (p. 443). Actually, however, the addition is quite permissible. It is correct that the additive property of chi-square assumes that the individual chi-squares are independent. This assumption correctly characterizes the theoretical model against which the observed frequencies are compared. If the probability corresponding to the sum of the chi-squares is small (less than five per cent or less than one per cent) then we may infer either that an unusual event has occurred, or that the hypothesis underlying the theoretical model is incorrect, or that some condition underlying the chi-square test was violated.

---

[1] Lewis, Don, & Burke, C. J. The use and misuse of the chi-square test. *Psychol. Bull.* 1949, **46**, 433–489.

Although the authors deny the applicability of the additive property of chi-square to the particular example, they do affirm the applicability of chi-square to the first toss of the coin because "on a single toss, the guess of each student was independent of the guess of the other students" (p. 443). Following the thinking of the authors, the chi-square text should not be even applicable to the first toss because the guess of one student is not necessarily independent of the guesses of the other students. There could be a general cultural bias in favor of heads on the first toss. In addition, in the same sense in which the second guess of a student may be related to his first, the first guess may be related to some psychological tendency of the student.

5. On pp. 478–480 the authors present an example in detail in order to clarify the meaning of "independence between measures." The discussion of this example is quite confusing.

Each of 240 students was asked to guess the fall of four successive tosses of a coin. The authors use the binomial formula in order to determine the probability distribution of the number of hits. It is assumed that the response of a subject is just as likely to be "head" as "tail." Since the probability of a hit is $\frac{1}{2}$ the theoretical number of 0, 1, 2, 3, 4 hits can be calculated. This distribution is then compared with the observed distribution of hits. The value of chi-square in this problem is equal to 43.351, with four degrees of freedom (see Table XXV, p. 479). Since the probability associated with this value of chi-square indicates the occurrence of a very unusual event, one would expect that the authors would reject the hypothesis. Since the original hypothesis was that the probability of a hit was equal to $\frac{1}{2}$, an alternative hypothesis would be one for which the probability of a hit is greater than $\frac{1}{2}$. The practical inference is that either the individuals responded with some knowledge of the actual fall of the coin or that the average response patterns of the subjects happened to coincide with the fixed pattern of the four coins (*H T T H*).

---

[2] Authors' italics.

Rather than reject the hypothesis or question the associated conditions, the authors come to the remarkable conclusion that "The absence of independence and the consequent inability to obtain unequivocal theoretical (chance) frequencies made this application of the chi-square test a meaningless procedure" (p. 480). By the phrase "absence of independence" the authors mean that there is some psychological linkage between the responses of a subject to the four successive tosess of a coin. Thus it is that the authors reject the application of the binomial formula in this particular problem (p. 479). It should be noted, however, that the application of the binomial formula assumes independence in a statistical sense, viz., that the joint probability of two random events is equal to the product of the respective probabilities of the two events. The binomial formula can be correctly applied to the problem on hand because the probability of getting a hit is not affected by the psychological tendencies of the subject. Of course, since the pattern of the four coins was fixed, *H T T H*, it could happen that psychological tendencies of the subject may produce a disproportionate number of hits. To obviate this factor it is necessary to toss the four coins for each trial. The thing which is at fault, if there is any fault, in this application is not the binomial formula but the structure of the experiment.

6. In a criticism of a problem cited from the literature, Lewis and Burke fail to mention the significant point that the sum of the probabilities is greater than unity (viz., 1.7), whereas the sum of the probabilities must be equal to unity (Table XXVI, p. 481). Moreover, the authors' own alternative analysis of the same problem contains the same error, the sum of the probabilities being somewhat larger than 1.8 (Table XXVII, p. 482).

Received January 3, 1950.

ON "THE USE AND MISUSE OF THE CHI-SQUARE TEST"—THE CASE OF THE 2×2 CONTINGENCY TABLE

ALLEN L. EDWARDS
The University of Washington

In their recent article in this journal, Lewis and Burke (**7**) provide an excellent guide for the research worker in psychology in terms of what to do and what not to do in applying the χ^2 test of significance. It is, of course, obvious that there are conditions which will invalidate any test of significance. In the case of the χ^2 test, authorities tend to agree that one of these conditions is the presence of small theoretical frequencies. That Lewis and Burke are seriously concerned about what they believe to be violations of this principle by psychologists is indicated by the fact that warnings against the use of small theoretical frequencies appear through their paper. For example:

> The use of "small theoretical frequencies" is listed as one of the "principal sources of error" in applications of the χ^2 test by psychologists in articles published in the *Journal of Experimental Psychology* over a three-year period (**7**, pp. 433–434).
>
> "The commonest weakness in applications of the χ^2 test to contingency tables is the use of extremely small theoretical frequencies" (**7**, p. 460).
>
> If we have but a single *df*, as in the case of the 2×2 contingency table, "the use of theoretical frequencies of less than 10 should be strictly avoided" (**7**, p. 460).
>
> "Whenever small theoretical frequencies enter into calculations of χ^2, the experimenter has no sound basis either for accepting or rejecting a hypothesis except when the value (of χ^2) is quite extreme" (**7**, p. 462).
>
> Several applications of the χ^2 test by Lewis and Franklin (**9**) are judged not acceptable because "they involved extremely small theoretical frequencies" (**7**, p. 433).

In this paper we shall be concerned only with applications of the χ^2 test to the 2×2 contingency table. More specifically, we shall be concerned with what constitutes a small theoretical frequency in this application.

1. There is no quarrel with the idea that a large number of well-controlled observations is more satisfying than a small number of well-controlled observations. However, if only a limited number of well-controlled observations can be made, this may be more satisfying than a large number of poorly-controlled observations. At the same time, with other conditions remaining constant, a small number of observations will yield smaller theoretical frequencies than a large number of observations. Let us grant, then, that a large *N* is a good thing, if observations are well-controlled, and that we like to have it when possible. This, in turn, means that we like to avoid small theoretical frequencies. There is no disagreement on this point.

Disagreement does occur when we ask the pertinent question: How large must a theoretical frequency be before it is considered not small? Or, to turn the question around: How small may a theoretical frequency be before it is considered not large? The problem of adequately defining smallness, in the case of the χ^2 test, is, as Yule and Kendall (**12**, p. 422) have pointed out, not a simple one.

Cramér (**1**), according to Lewis and Burke, "firmly recommends a minimal value of 10" (**7**, p. 487), and Lewis and Burke are quite explicit in stating that in their opinion: "A value of 5 is believed to be too low" (**7**, p. 462). Yule and Kendall (**12**, p. 422) take a more moderate view. They regard 5 as a minimal value, but add that 10 is better. Hoel (**5**, p. 191) suggests that 5 is satisfactory if the number of cells or categories is equal to or greater than 5. But this is not the case with the 2x2 contingency table and, in this instance, Hoel suggests that it is better to have theoretical frequencies "somewhat greater than 5" (**5**, p. 191). Fisher (**3**, p. 87, 97), on the other hand, seems content with the rule of 5.

The position taken in this paper is that psychologists (and other research workers) are not necessarily "misusing" the χ^2 test when they apply it to the 2x2 contingency table when a theoretical cell frequency is as low as 5.

2. The exact probability for any set of frequencies in a 2x2 table can be obtained by direct methods of calculation as given by Fisher (**3**, pp. 100–102). At the same time, it is important to emphasize, as Fisher has, that the χ^2 test applied to the same data "can only be of approximate accuracy" (**3**, p. 97), and that its usefulness lies in "the comparative simplicity of the calculations" (**3**, p. 100).

The "approximate accuracy" of the χ^2 test can be judged by the extent to which the probability associated with the obtained value of χ^2 is in accord with that obtained by the direct method. Agreement between the two values of P is likely to be best when theoretical frequencies are not small. The reasons for this are well stated by Lewis and Burke (**7**, pp. 439–440). This does not mean, however, that χ^2 is not useful for its "approximate accuracy" in estimating the exact probabilities in 2x2 tables involving theoretical frequencies as small as 5. Nor is the application of the χ^2 test in this instance a "misuse" of χ^2.

Consistently rejecting a null hypothesis when P is .05 or less and accepting the hypothesis when P is greater than .05 is a practice followed by many research workers. It is suggested here that the research worker will not be led badly astray in evaluating the 2x2 contingency table with theoretical frequencies as small as 5 (1) if he consistently follows the P equal to or less than .05 standard (or some other standard which is acceptable); (2) if his attitude toward the null hypothesis is based upon the value of χ^2 corrected for continuity; and (3) if he calculates the exact probabilities for those cases where the probability value obtained by the χ^2 test is of borderline significance.

In practice, as Fisher (**3**, p. 83) states with his usual common sense attitude, our interest is not primarily in the exact value of P for any given hypothesis, but rather in whether or not the hypothesis is open to suspicion.

3. Let us examine the two sets of data given by Lewis and Burke involving a 2x2 table with small theoretical frequencies. Table 1, given below, is their Table XVI, and is based upon the data of Lewis (8) and Lewis and Franklin (**9**). With respect to this table, Lewis and Burke state that "all four of the theoretical frequencies . . . are too small to warrant an application of the χ^2 test" (7, p. 461).

TABLE 1

CLASSIFICATION OF SUBJECTS WORKING UNDER DIFFERENT CONDITIONS ON THE BASIS OF THE RATIO (*RI/RC*) OF THE NUMBER OF INTERRUPTED TASKS RECALLED TO THE NUMBER OF COMPLETED TASKS RECALLED*

RI/RC	*Individual Work Condition*	*Cooperative Work Condition*	*Total*
Greater than 0.67	6	12	18
Less than 0.67	6	2	8
Total	12	14	26

* This table corresponds to Table XVI in the Lewis and Burke article (7, p. 461). The data presented are based upon experiments by Lewis (8) and Lewis and Franklin (9).

Introducing Yates' correction for continuity in one of the standard computational formulas for χ^2, we have

$$\chi^2 = \frac{N\left(bc - ad - \frac{N}{2}\right)^2}{(a+b)(a+c)(b+d)(c+d)}, \qquad [1]$$

where the letters a, b, c, and d correspond to the cell frequencies and N represents the total number of observations. Calculating χ^2 for the data of Table 1, we obtain a value of 2.374. And since χ, the square root of χ^2 with 1 *df*, is distributed as a normal deviate z, we find from the table of the normal curve that when $z = 1.54$, P is .0618. The corresponding value of P for χ^2 will be (2) (.0618) = .1236.[1] What is the "approximate accuracy" of this value in terms of the probability obtained by direct methods?

Assuming the constancy of the marginal totals, as is assumed in calculating χ^2 also, the probability of any observed set of frequencies is given by the product of the factorials of the four marginal totals, divided by the product of the factorials of the grand total and the four cell entries (**3**, pp. 100–102). In terms of a formula[2]

---

[1] The reason why we take $2P$ for the probability associated with χ^2 is discussed in Edwards (2) and Goulden (4).

$$P = \left(\frac{(a+b)!(a+c)!(b+d)!(c+d)!}{N!}\right)\left(\frac{1}{a!b!c!d!}\right). \qquad [2]$$

The desired probability, however, involves not only the arrangement of frequencies as given in Table 1, but all other possible arrangements (deviations) which are more extreme. Thus we need the probabilities for the following arrangements:

$$\begin{matrix} 6 & 12 \\ 6 & 2 \end{matrix} \qquad\qquad \begin{matrix} 5 & 13 \\ 7 & 1 \end{matrix} \qquad\qquad \begin{matrix} 4 & 14 \\ 8 & 0 \end{matrix}$$

The probability desired will then be the sum of the probabilities for the three arrangements. Direct calculation shows this to be .0612, a value which may be compared with the probability of 0.618 for χ. We also have (2)(.0162) =.1224 with which we may compare the probability of .1236 for χ^2. Not only would no conclusions concerning significance be changed, if χ^2 is calculated with the correction for continuity, but the "approximate accuracy" of the χ^2 test seems quite good in this particular instance.

Table 2 reproduces Kuenne's (**6**) data as given by Lewis and Burke in their Table XVII and for which they are also concerned about the application of the χ^2 test because of small theoretical frequencies. From formula [1] we find χ^2 to be 13.1585, and χ to be equal to 3.627. From the table of the normal curve, in the manner previously described, we find that the value of P corresponding to χ is .0001433, and for χ^2 we have (2) (.0001433) = .0002866. Direct calculation, in the manner described earlier, shows that the corresponding values of P are .0000985 and (2) (.0000985) = .0001970, respectively. Again no conclusions concerning significance would be changed, regardless of whether the data

TABLE 2

CLASSIFICATION OF SUBJECTS IN TWO AGE GROUPS ON THE BASIS OF THOSE SHOWING TRANSPOSITION AND NON-TRANSPOSITION*

Age Groups	*Transposition*	*Non-Transposition*	*Total*
3–4 years	3	15	18
5–6 years	20	6	26
Total	23	21	44

* This table corresponds to Table XVII in the Lewis and Burke article (**7**, p. 462). The data are based upon an experiment by Kuenne (**6**).

---

[2] A table of the logarithms of factorials such as may be found in Pearson's tables (**10**) or the *Mathematical Tables from Handbook of Chemistry and Physics* (**13**) facilitates the necessary calculations.

are evaluated by means of χ^2 or by the direct method. As to the "approximate accuracy" of the χ^2 test, it may be observed that the discrepancy between the two values of P for χ^2 and for the direct method is .0000896.

A few additional examples taken from Goulden (**4**) and Yates (**11**) illustrate the degree of "approximate accuracy" which χ^2 may give, even when theoretical frequencies are smaller than the value of 5. In these examples the probability obtained by the direct method has been given by the source from which the data are taken. χ^2 has in all cases been corrected for continuity. For simplicity, the value of P is given for χ and as obtained by the direct method. The value P for χ^2 and the corresponding value for the direct method may be obtained by doubling the reported values. Examples 1, 2, and 3 are real in that they are based upon actual data, and Example 4 is simply an exercise taken from Goulden.

Example 1—Hellman's data cited by Yates (**11**).

	Normal Teeth	*Malocclusion*	*Total*
Breast-fed	4	16	20
Bottle-fed	1	21	22
Total	5	37	42

$\chi = 1.068$: $P = .1427$
Direct method: $P = .1435$

Example 2—Grant's data cited by Goulden (**4**).

	Group Blood 0	*Blood Group Not 0*	*Total*
Fond du lac Indians	18	11	29
Chipewyan Indians	13	1	14
Total	31	12	43

$\chi = 1.75$: $P = .0401$
Direct method: $P = .0349$

Example 3—Mainland's data on the position of polar bodies in the ova of the ferret cited by Goulden (**4**).

	Similar	*Different*	*Total*
10μ apart	5	1	6
More than 10μ apart	1	6	7
Total	6	7	13

$\chi = 1.93$: $P = .0268$
Direct method: $P = .025$

Example 4—Exercise given in Goulden (**4**).

	Recovered	*Died*	*Total*
Animals inoculated	7	3	10
Animals not inoculated	3	9	12
Total	10	12	22

$\chi = 1.68$: $P = .0465$
Direct method: $P = .0456$

As will be apparent, each of the examples cited involves theoretical frequencies of less than 5, yet the degree of "approximate accuracy" of the χ^2 test seems satisfactory. We should not expect, however, that all 2x2 tables with theoretical frequencies as small as these will yield P's, when tested by means of the χ^2 test, which are as close to those obtained by the direct method.

4. It is to be emphasized that there is no magic involved in the rule of 5 as a minimal theoretical frequency—any more than there is in the rule of 10. It will be true, however, that if we calculate the value of P directly whenever the theoretical frequency is under 10, we shall be involved in roughly about twice as much labor as would be the case if we accept 5 as the minimal theoretical frequency.

The procedure suggested here is to apply the rule of 5 and to calculate χ^2, keeping in mind the correction for continuity. If the obtained value of χ^2, in such cases, is of borderline significance, then calculate the value of P by the direct method. If theoretical frequencies of less than 5 are involved, the desired probability may be found, not too inconveniently, by the direct method.

BIBLIOGRAPHY

1. Cramér, H. *Mathematical methods of statistics*. Princeton: Princeton Univ. Press, 1946.
2. Edwards, A. L. *Experimental design in psychological research*. New York: Rinehart, 1950.
3. Fisher, R. A. *Statistical methods for research workers* (6th Ed.). Edinburgh: Oliver and Boyd, 1936.
4. Goulden, C. H. *Methods of statistical analysis*. New York: Wiley, 1939.
5. Hoel, P. G. *Introduction to mathematical statistics*. New York: Wiley, 1947.
6. Kuenne, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. exp. Psychol.*, 1946, **36**, 471–490.
7. Lewis, D., & Burke, C. J. The use and misuse of the chi-square test. *Psychol. Bull.*, 1949, **46**, 433–489.
8. Lewis, H. B. An experimental study of the role of the ego in work. I. The role of the ego in cooperative work. *J. exp. Psychol.*, 1944, **34**, 113–126.
9. Lewis, H. B., & Franklin, M. An experimental study of the role of the ego in work. II. The significance of task-orientation in work. *J. exp. Psychol.*, 1944, **34**, 195–215.
10. Pearson, K. *Tables for statisticians and biometricians*. Cambridge: Cambridge Univ. Press, 1914.
11. Yates, F. Contingency tables involving small numbers and the χ^2 test. *Suppl. J. Royal stat. Soc.*, 1934, **1**, 217–235.
12. Yule, G. U., & Kendall, M. G. *An introduction to the theory of statistics*. (13th Ed.). London: Griffin, 1947.
13. *Mathematical tables from handbook of chemistry and physics* (7th Ed.) Cleveland: Chemical Rubber Publishing Co., 1941.

Received December 1, 1949.

FURTHER DISCUSSION OF THE USE AND MISUSE OF THE CHI-SQUARE TEST

DON LEWIS
State University of Iowa

AND

C. J. BURKE
Indiana University

The articles by Peters (5), Pastore (4), and Edwards (2) objecting in one way or another to our paper on chi-square (3), indicate either that the paper was not carefully studied or that our exposition was less clear than we had supposed. Since Peters and Pastore raise several points in common, we are forced to conclude that the fault was ours to a considerable extent. A clarification of the points in question will be attempted.

We should perhaps say, parenthetically, that we are confident we could have written a clear, concise paper in mathematical language which would have left little or no room for disagreement between Peters, Pastore, and Edwards, on the one hand, and ourselves, on the other. Because the paper was aimed at the average user of statistics among psychologists, we rejected a mathematical type of exposition in favor of proceeding largely by example. We are disappointed but not surprised to discover that our treatment has been found obscure in a few places.

REPLY TO PETERS

Before the writing of the paper was begun, we carefully surveyed the basic literature on chi-square, including several papers by Pearson, but could see no good reason for basing our discussion on the earliest proofs. Full credit is due Pearson for originating the chi-square test. His proofs are adequate for establishing the general validity of the statistic. However, advancements have been made in the mathematical methods applied to statistical problems; mathematicians have been able to improve upon many of the early proofs. Such recent treatises as Cramér's *Mathematical Methods of Statistics* (1) provide a more complete background for the χ^2 distribution function and for the tests associated with it, and they serve as a suitable point of departure for any discussion of χ^2. In making this statement, we intend no disparagement of Pearson. He falls easily among the four or five most eminent statisticians. It is a commonplace in science that later investigators, often persons of lesser ability, using the insights of a great innovator, frequently improve upon his work.

The specific points made by Peters will be discussed in the order in which he presents them.

1. *That we claim chi-square applies fundamentally only to frequencies.* An examination of our equations [1], [2], and [8] and of the comments in footnote 8, page 439, and on page 487 will show that we do no such thing. Our equation [8] is the exact equivalent of Peter's equation [iii], except for the error in the exponent of χ^2 which Peters and Pastore both mention. Several persons have written since November calling our attention to this error, and we discovered it ourselves soon after it appeared in cold ineradicable print. We will be greatly indebted to anyone who can provide a foolproof method of avoiding errors of this kind.

What we do claim is that the formula

$$\chi^2 = \sum \frac{(F_o - F_t)^2}{F_t} \qquad [10]$$[1]

can be legitimately applied only to raw frequency data. Applications involving frequencies are the ones of greatest interest to most psychologists, so it seemed appropriate to emphasize these applications and to enumerate certain fundamental assumptions and restrictions through a discussion of the binomial and multinomial distribution functions. We are well aware of the fact (stated in our footnote 10) that the chi-square distribution function may be derived from the joint normal distribution function in *n* variables, and aware of the additional fact (stated several times in the paper) that the chi-square statistic plays an important role in the study of variances.

2 *and* 3. *That the sum of the theoretical frequencies need not equal the sum of the observed frequencies, and that the frequencies of non-occurrence need not always enter into the calculations.* These two points can best be discussed together, as they are intimately related.

Cramér (1, p. 426 ff.) rigorously proves an underlying theorem on chi-square and indicates how the proof of a generalization of the theorem would be handled. The generalized theorem is broad enough to include as special cases all applications of chi-square which are based on formula [10]. It encompasses the tests of fixed probabilities, tests of goodness of fit, and tests of independence, and shows clearly the conditions which must prevail for other tests of the same general kind. It includes the case covered by Pearson as well as the new derivation requested by Peters in his third footnote, and establishes the fact that *the number of degrees of freedom is the same in both cases*. In Cramér's proof, it is necessary to assume that the frequency associated with every possible outcome of the experiment is used and that the sum of the theoretical frequencies is made equal to the sum of the observed frequencies. It should be emphasized that these restricting assumptions hold only for applications of chi-square based on formula [10].

---

[1] The formula carries this number in the original paper.

Peters states that he and Van Voorhis erred in one of their examples (6, Table XXXV) by using his formula[vi] instead of [v]. This is simply an alternative way of wording one of our objections. We show in our paper (3, pp. 437–438) that using his formula [v] is exactly equivalent to using our formula [10] and taking account of frequencies of non-occurrence after the sums of the observed and theoretical frequencies have been equated.[2] What Peters does, in effect, is to accept our criticism but insist that it be stated in another way.

Peters maintains that the application of chi-square which he and Van Voorhis made in relation to numbers of naturalized citizens among several nationality groups is a correct one. Since a careful examination of this application is instructive, we give in Table I, Part A, observed frequencies (of occurrence and non-occurrence of naturalization) adapted from Peters and Van Voorhis' Table XXXVI. In Part B of the table are the theoretical frequencies as calculated by Peters and Van Voorhis on the assumption of homogeneity of nationality groups with respect to the proportion naturalized. Note that the marginal totals in the two parts of the table are identical. The theorem given by Cramér provides unimpeachable proof that the quantity obtained by using formula [10] and summing over all eight categories will have a distribution approximating the χ^2 distribution with 3 *df* provided that the hypothesis of homogeneity is true. In this example, as can be readily verified, the value of χ^2 based on the eight categories is 29.72. Peters and Van Voorhis used formula [10] also, but excluded the "remainder" categories and

TABLE I

OBSERVED AND THEORETICAL FREQUENCIES FROM TABLE XXXVI IN PETERS AND VAN VOORHIS (6, p. 414)

A. OBSERVED FREQUENCIES

	Italians	*Russians*	*Polish*	*Others*	*Totals*
Number Naturalized	161	82	20	32	295
Remainder	205	34	32	22	293
Totals	366	116	52	54	588

B. THEORETICAL FREQUENCIES

	Italians	*Russians*	*Polish*	*Others*	*Totals*
Number Naturalized	183	58	26	28	295
Remainder	183	58	26	26	293
Totals	366	116	52	54	588

---

[2] In this particular case, by using [v] Peters obtains a correct value of 12, as he reports in his second footnote (5, p. 332). Using equation [10], we obtained the correct value of 12 and so stated in our fourth footnote (3, p. 435). This agreement is not a coincidence.

summed over only the top four categories, to obtain a value of 14.52. Since their computations included only half of the categories, they would obtain for any similar sample a value considerably smaller than we would obtain, regardless of the truth or falsity of the hypothesis. The value computed from all eight categories is known to have, approximately, the χ^2 distribution with 3 *df*, when the hypothesis is true. It follows, of course, that the quantity they computed cannot have the distribution they attribute to it because it is systematically too small.[3] Thus, there is no logic for interpreting their statistic by means of the χ^2 distribution with 3 *df*. The fact that the decision reached in this particular case would be the same for both computational methods is completely beside the point.

Peters seems to imply that there are no differences between an application of this kind and the tests of goodness of fit for ordinary distribution functions. Where, we ask, in testing the goodness of fit of a normal curve, do categories of "success" and "failure" ever exist over the same class interval? The difference between the applications is really quite profound.

4. *That no difficulties arise in applying chi-square in situations where the presence of individual differences may bring about a lack of independence.* Here Peters misunderstands what we mean by *independence*, a misunderstanding that often arises because the term is used with at least four distinct meanings in connection with chi-square tests. We have in mind the well-known distinction between the Bernoulli-De Moivre theorems on the large sample binomial distribution (based on constant probabilities) and the Poisson generalization of these theorems (including variable and linked probabilities). A somewhat extreme, but simple, example will clarify our objection to the dice-throwing illustration used by Peters and Van Voorhis.

Suppose that we have 20 coins which are of three kinds. With two of them, the probability of obtaining either a head or a tail is $\frac{1}{2}$; but nine of the coins have heads on both sides while the remaining nine have tails on both sides. If we toss the 20 coins together we will observe either 9, 10 or 11 heads, the remainder being tails. From such a toss, the value of χ^2 computed with formula [10] with theoretical frequencies of 10 must be either 0.0 or 0.2, with a probability of $\frac{1}{2}$ associated with each. If we were to toss the coins five times and combine the results as Peters and Van Voorhis did in their dice-throwing illustration, we would obtain for the computed quantity the following exact distribution:

χ^2—0.0	0.2	0.4	0.6	0.8	1.0
P—1/32	5/32	10/32	10/32	5/32	1/32

If a quantity having this distribution were interpreted in accordance with the χ^2 distribution, the hypothesis that the coins are unbiased would never be rejected since, with 5 *df*, a χ^2 value larger than 1.0 has a probability of occurrence of over 0.95. And yet, we know that the coins are biased.

---

[3] On the same grounds, we reaffirm our statement that Grant and Norris erred in omitting frequencies of non-occurrence.

Too much emphasis should not be placed on the incorrectness of adding χ^2 values in a situation of this kind. The results would not be much better if the hundred turns (20 coins each tossed 5 times) were treated as a unit. The use of formula [10] with the pooled results, the theoretical frequencies now being 50, would yield a quantity having the following exact distribution:

χ^2—0.0	0.04	0.16	0.36	0.64	1.0
P—126/512	210/512	120/512	45/512	10/512	1/512

Again, the hypothesis that the coins are unbiased would never be rejected through the use of the χ^2 distribution because, with 1 *df*, the probability of obtaining a value of χ^2 larger than 1.0 by chance is over 0.3.

In order to make a meaningful study of the coins, we would be obliged to specify (if we could) some event which had a uniform and not a varying probability of occurrence. With a population of coins of the three specified types, we might test the hypothesis that the probability of selecting a coin at random, tossing it, and observing a head is $\frac{1}{2}$. To do this, we would need to select different coins for each repetition of the experiment (a procedure which apparently surprises Pastore).

We would not object quite so strongly to the dice-throwing illustration of Peters and Van Voorhis if there were not a simple modification in procedure which eliminates the difficulty. If each die were tossed a sufficient number of times to permit the computation of separate values of χ^2 to represent the "behavior" of the separate dice and the χ^2 values were then summed over the set of dice, the deviations due to individual differences would appear. In this design, there could be no compensating effects from one die to another. The same general procedure could be used for generalizing to a larger collection of coins.

The objection may be raised that our coin example is too bizarre to be taken seriously. (Who ever heard of a coin with heads or tails on both sides?) Our answer is that, in any situation where individual differences in the underlying probabilities exist, the same effect is present although usually not so conspicuously. We are inclined to believe that individual differences among coins and dice may actually contribute to their ways of turning. We insist that with subjects in psychological experiments, individual differences cannot be ignored.

Peters finds us picayune in our discussion of one of his examples, and his attitude is not entirely unwarranted. However, for better or worse, we are forced to judge his book partly by the uses our students and colleagues make of it. Some psychologists are prone to take any example in an authoritative text as a model of experimental design and to copy it in their own work. The text by Peters and Van Voorhis is widely used, and justifiably so; in many respects, most notably in its exposition of mathematical backgrounds, it is an admirable book. This wide use imposes certain obligations on the authors. We suspect that they may be partly responsible for experimental designs which they would certainly not sanction. This is admittedly a poor way to judge books, but in the field of statistics we have no other recourse. Because of the unexpected uses to which examples may be put, we believe that correctness should never be sacrificed to pedagogical simplicity.

5. *That we erred in not attributing to Karl Pearson the notion that a quantity computed according to fixed rules can have two different distributions depending on the choice of the experimenter.* We readily concede that Pearson discusses this point in almost the same way that Peters and Van Voorhis do. And yet, we find a certain ambivalence in his position. At times, he writes as if he regarded his recommendations as approximations which are acceptable only on empirical grounds. It comes down to a matter of interpretation, and we doubt that anything can be proved one way or the other. We retract the statement to which Peters objects as being too firmly worded and admit that his interpretation of Pearson's views may be the correct one. However, the retraction does not alter the fact that the number of degrees of freedom for χ^2 in situations of the type discussed by Pearson and Peters is unaffected by the way the investigator phrases the hypothesis under test.

6. *That the χ^2 test of linearity of regression is adequate in some circumstances, but can be replaced by a better test.* This is precisely the view taken in our paper. We might have attributed the χ^2 linearity test to Fisher, as Peter says we should have done, but it was not our aim to designate the originator of every technique we mentioned. For our purposes, the best points of departure were the formulas and recommendations in the texts and reference works most widely used by psychologists.

We share many of the views expressed by Peters in his closing paragraphs but are less sanguine than he concerning the likelihood of finding suitable replacements for χ^2.

REPLY TO PASTORE

Several of Pastore's criticisms have been answered either directly or indirectly in our reply to Peters. Those which remain will now be discussed, with the numbering kept the same as his.

2. He contends that we fail to offer a better method for handling the dice-throwing problem of Peters and Van Voorhis. We actually present a correct method (**3**, pp. 446–447) indicating that separate values of χ^2 should be obtained for the individual dice and that these

values could then be summed. This procedure is without flaws. The method proposed by Pastore is also a correct one and would avoid the difficulties posed by the possibility of individual differences, but a large number of throws (much larger than 14) would be required. Also, because there are 13 categories, there would be 12 *df* instead of 11, as stated by Pastore—provided, of course, that enough throws were made to yield theoretical frequencies of acceptable size in all categories.[4] With the technique we suggest, the particular dice (if any) which were biased could be spotted; with Pastore's technique, this would not be possible.

4. Pastore maintains that values of χ^2 can be combined even when they are known to be interdependent, provided that there is independence in the underlying model. It is made quite clear in our paper (**3**, p. 442 ff.) that we are discussing a case in which linkages between guesses on separate trials are empirically known to be present. Frankly, we are rather surprised at being asked to defend the thesis that a test should not be based on a demonstrably false assumption, and wonder why anyone would wish to use a model which is already known to be inappropriate. However, we shall present an artificial example of a rather extreme type to clarify the major issue.

Suppose that we run a coin-guessing experiment of the kind described in our original paper, but someone instructs the subjects to alternate their guesses throughout the experiment. We happen to select a group of 100 subjects such that, on the first guess, 55 responses are heads and 45 are tails. The resulting value of χ^2 is 1. On the second guess, the responses are 45 heads and 55 tails, and the value of χ^2 is again 1. After 20 guesses, we would have a composite χ^2 value of 20, with 20 *df*. Since a value larger than this would have a probability greater than 0.4 of occurring by chance, we would retain the hypothesis that each guess was mediated entirely by chance factors. But in this situation, the only random event is the first guess. We wonder if Pastore would wish to consider this set of guesses as a purely chance affair throughout.

If the experiment were repeated under the same conditions, 43 guesses of heads and 57 guesses of tails might be obtained on the first trial. The resulting value of χ^2 would be 1.89. Larger values than this have a probability roughly equal to 0.15 associated with them. The difference between 1.89 and 1.00 is not large and little importance would be attached to it. After 20 guesses, however, we would have a composite value of 37.8, significant at the 1% level. Thus, under these conditions, first-trial differences which are unimportant can lead, after summation, to rather extreme differences in interpretation.

---

[4] The probability of 12 aces on a single throw would be about .000000000459, so it would take around 20 billion throws to give a theoretical frequency of 10 in that category.

Whenever interdependencies exist, effects of the type just illustrated will be present. Pastore's extension of our argument to the first guess is clearly illegitimate. Any deviation of the frequencies from the theoretical values to be tested can be unambiguously interpreted and no difficulties arise. Also, if a different sample of subjects is used on each guess in obtaining a composite value of χ^2, there are no interdependencies to complicate the result.

5. It is asserted that, in one of our examples illustrating indeterminate theoretical frequencies, the test which we reject is a correct one. Since a detailed discussion of this point turns upon interdependencies of the kind already treated in detail and would merely be a repetition of an argument previously given, we shall confine ourselves to a few general remarks. (If one considers what would happen with a group of subjects who were systematically alternating their guesses, the difficulties become quite clear.)

Pastore may be justified in saying that the obtained value of $\chi^2(=43.351$, with 4 *df*) is large enough to warrant a rejection of the hypothesis without further question. But here again, he shows a willingness to use a model that is known to be inapplicable. What can be gained from such a procedure? If the hypothesis is rejected, nothing is gained because it was already known to be false. If it is not rejected, it still cannot be retained in view of its known falseness.

Pastore writes as if a statistic can be correct for an experiment even though the structure of the experiment is faulty. As far as we are concerned, the experimental design and the statistical treatment constitute a unit and must be treated as such. Incidentally, we should like to learn Pastore's technique for testing the hypothesis that a probability is greater than $\frac{1}{2}$.

6. Pastore objects because we include (**3**, pp. 481–482) two sets of probabilities which sum to more than unity. Considered in their proper setting, they are obviously conditional probabilities. The sum of the corrected probabilities is unity in each case, since the sum of the observed frequencies is equal to the sum of the theoretical frequencies. Perhaps this point can be made clear by an example which is a close parallel to the case under consideration.

Suppose that someone offers us a gift which is hidden under one of three upturned cups. The gift is to be ours if we can guess the cup under which it lies. The probability is $\frac{1}{3}$ that we will guess right. Our first guess is wrong, but we are invited to try again. The cups and gift remain as they were. The probability of a correct guess is now $\frac{1}{2}$. Following our second incorrect guess, a third one is proffered. Our memory is sufficiently good—under ordinary circumstances—to make the third probability 1.0. The sum of this set of probabilities is 1.83.

REPLY TO EDWARDS

Edwards' paper (2) seems less a criticism of ours than an attempted advance based on a new collection of evidence; and we doubt that there are any serious disagreements between him and us. Our paper, although strongly recommending a minimal value of 10 for theoretical frequencies, allows for smaller values in the 2×2 table and proposes the use of Fisher's exact treatment whenever they occur. Edwards argues for frequencies as low as five and for the use of χ^2 as a sort of screening device and the use of the exact treatment as a check. The difference in viewpoint is not very great.

We may have overstated our case. At least Edwards' examples are quite convincing as far as they go, and his recommendations seem reasonable.[5] Nevertheless, we are inclined to take a position of "wait-and-see." In many cases, as shown by Edwards' examples, chi-square tests based on small theoretical frequencies are not far off, and there appears to be the possibility of proving that such tests will never be far off. Until such proof is forthcoming, however, the possibility must be accepted that combinations of numbers in fourfold tables may exist for which the divergences could not be neglected. In this connection, we will be quite happy to see final proof that we are wrong.

As a final word, we wish to express the sincere hope that everything we have written on chi-square will (to borrow a phrase from Peters) "be read with critical alertness, not taken offhand as gospel."

BIBLIOGRAPHY

1. CRAMÉR, H. *Mathematical methods of statistics*. Princeton: Princeton Univ. Press, 1946.
2. EDWARDS, A. L. On "The use and misuse of the chi-square test"—The case of the 2×2 contingency table. *Psychol. Bull.*, 1950, **47**, 341–346.
3. LEWIS, D., & BURKE, C. J. The use and misuse of the chi-square test. *Psychol. Bull.*, 1949, **46**, 433–489.
4. PASTORE, N. Some comments on "The use and misuse of the chi-square test." *Psychol. Bull.*, 1950, **47**, 338–340.
5. PETERS, C. C. The misuse of chi-square—A reply to Lewis and Burke. *Psychol. Bull.*, 1950, **47**, 331–337.
6. PETERS, C. C., & VAN VOORHIS, W. R. *Statistical procedures and their mathematical bases*. New York: McGraw-Hill, 1940.

Received May 1, 1950.

---

[5] With reference to his reworking of the frequency data obtained by Lewis, Lewis and Franklin, and Kuenne, we disavow in our paper any criticism of the experimental conclusions reached by these investigators.

Dashiell, J. F. *Fundamentals of general psychology.* (3rd Ed.) Boston: Houghton Mifflin, 1949. Pp. x+690. $4.00.

It is gratifying to have a new edition of Professor Dashiell's well known and respected text, *Fundamentals of General Psychology*, in a new format and written in a style which should hold student interest without loss of scientific soundness. The general outlook of the book has changed little and remains essentially eclectic with a leaning toward the language of stimulus-response psychology. However, this edition reflects the increased interest in dynamic problems which is evident among academic psychologists. A few examples will illustrate this point. In 1937 Freud's familiar defense mechanisms were discussed but without reference to Freud. In the new edition Freud is mentioned in a number of places, though somewhat reluctantly and with a warning to the student. The Freudian concept of personality structure is presented and the notion of libido is mentioned briefly in the section on sex drive. In addition, Dashiell now makes use of the ego as an organizing factor in motivation and finds the concepts of identification and levels of aspiration to be useful in this connection.

The student is presented with a large amount of original experimental material, and in general one cannot quarrel with the accuracy of the reporting. It is clear and to the point. We must however, question at least one interpretation. Dashiell, in common with a great many writers, talks about determining the percentage of a person's achievement which is due to heredity and to environment, and suggests that the results indicate the contribution to be "fifty-fifty." There is of course *no* method available for determining the percentage of *a person's* achievement or development which is due to heredity or to environment. The fact is we can know only that a certain percentage of *variation* between individuals in a group is due to the *differences* in these factors.

It is always important to ask how a text handles information about the tools and techniques of psychology, for the intelligent layman needs to develop critical skills in these areas if he is to be protected against half-truth. Dashiell discusses various kinds of observation and measurement in almost every chapter. However, despite the fact that considerable space is devoted to relatively complicated statistical concepts, such as correlation and critical ratios, one misses the explanation of what psychological measurements mean in logical terms. There is, for example, no reference to the limitations of relative measures, to the absence of absolute zero or of equal units. It is the reviewer's experience that these concepts are more easily understood and more usefully employed by the elementary student than is the arithmetic of correlation. Also missing is a description of the problem of the inferred variable. Dashiell discusses attitudes and sets and motives and other concepts of this sort which the psychologist must use in his science. He does not make clear to the student that such notions are different in a very important way from the more direct observations of stimulus and response, which are regarded as fundamental. The student of general psychology or of any science should become familiar with the fact that inferences of this kind are among the most important activities of the scientist, whether he be a physicist or a psychologist.

Textbooks vary in their degree of organization around systematic principles. The Dashiell text takes a middle course in this respect. The principles used are eclectic and not always definite. This is true, for example, of the attitude toward learning. Learning is "organismic adjustment." The author describes various learning processes, admits that the law of effect holds if it is looked at broadly enough, accepts contiguity, sees that learning may be facilitated by observation, and says that some day, "all partial, single-viewpoint theories will become successfully incorporated into one complete and completely satisfactory system of learning principles." He does not attempt to indicate what this system is likely to be, but it must be admitted that there is a real question whether such a commitment can be made at the elementary level without unwarranted dogmatism.

A final word should be said about the treatment of perception, for this illustrates better than most topics the fact that Dashiell retains the spirit of a stimulus-response approach in spite of his greater use of what is generally called dynamic material. The chapters on sensory functions, set, attention, and perceiving are widely separated. This results in what seems to the reviewer to be a lack of integration of the problems of sensitivity, judgment, perception and motivation. But for any author the words of his sentences must proceed serially and he cannot discuss everything simultaneously. This inevitably makes difficult the description of fluid and shifting and interrelated psychological processes which rarely fit the limitations of sentence structure.

HELEN PEAK.

Connecticut College.

O'KELLY, LAWRENCE I. *Introduction to psychopathology.* New York: Prentice-Hall, 1949. Pp. xxi+736. $4.50.

O'Kelly's justification for the above title can be gleaned from his statement that "An examination of current texts in 'abnormal psychology' shows that the selection of subject matter has changed over the years, and that such texts are, in fact, psychopathologies at the present time and are not the collection of curiosities that include such 'normal' abnormalities as sleep and hypnosis. The phrase 'abnormal psychology' is as awkward and inappropriate to our modern knowledge of maladaptation as the phrase 'abnormal anatomy' or 'abnormal physiology' would be as a title for texts in tissue pathology." In this book the author has made an able and comprehensive start in demonstrating the fruitfulness of viewing pathological phenomena as instances of broad psychobiological principles. The student reading this book is not likely to come away with the idea that different principles are necessary to explain different pathological conditions.

This text is divided into three parts. In the first part (Introduction) the major emphasis is on the definition of basic concepts. The reader is not plunged into a discussion of pathological phenomena but rather is given some broad principles which are necessary for understanding any aspect of behavior. In the second part (The Problems of Disordered Behavior) there are better and more extended discussions of anxiety and psychosomatics than will be found in most other texts. This part also has an adequate coverage of the psychoses and psychoneuroses. The third part (The Causes of Disordered Behavior) contains discussions of heredity, mental deficiency, organic factors, social factors, theories of psychopathology, and treatment. Throughout the book the author's training and work in comparative psychology are reflected in his lucid and comprehensive handling of the physiological aspects of the subject matter.

This reviewer would recommend this book for undergraduate courses with very few reservations. There are several shortcomings which should be kept in mind when this text is used in graduate courses. The discussion on theories of psychopathology is meagre; problems in treatment and prophylaxis are also given too little attention; test data continue to be viewed as the sole criterion of mental deficiency; and cross-cultural data are not given the discussion they deserve. It should be said that these shortcomings hold for all other texts which this reviewer has read. The virtues of this book far outweigh its shortcomings. On the assumption that in class the instructor does more than tell students what they already have read in their assignments, this book can be heartily recommended for graduate courses.

Yale University. SEYMOUR B. SARASON.

SALTER, ANDREW. *Conditioned reflex therapy—The direct approach to the reconstruction of personality.* New York: Creative Age Press, 1949. Pp. x+359. $3.75.

No psychologically informed person will accept this book, but neither will he be able to dismiss it lightly.

Here are some of the things that make the book vivid and arresting:

1. The author can write! His ability to turn a neat phrase and to express his meaning in terse, emphatic prose will be the despair of many an academician.

He knows how long to make a chapter and how to organize a book so as to sustain interest and be pedagogically effective.

2. The author is well read. Each chapter has a technical bibliography, and the book is seasoned with apposite literary allusions.

3. Whatever criticisms are leveled against the book, no one can fairly accuse the author of being clinically naive. He is manifestly experienced, hard-bitten. He often condenses into one or two sentences clinical inductions and insights to which more prolix writers would devote a paragraph or a page.

4. The author is right in maintaining that psychotherapy is a matter of *learning*. He is undaunted by medical pretensions, and unblushingly makes therapy a *psychological* enterprise.

Some of the patent weaknesses of the book are:

1. As the title—*Conditioned Reflex Therapy*—suggests, the author takes Pavlov as his principal authority. For him (the author), a neurotic is simply a person who in childhood has acquired the wrong ("inhibitory") conditioned reflexes. In therapy (by Salter) he acquires the right ("excitatory") ones. "The sickness is always inhibition, and the medicine is always excitation" (p. 300). Why old, inappropriate "conditioned reflexes" do not undergo spontaneous extinction (as Pavlovian theory would suggest) is a question which the author does not consider.

2. Explicitly, "learning" is equated to "conditioning." Yet the author repeatedly acknowledges a form of learning (trial-and-error, problem-solving) which is determined by motivation and reward. As an avowed reflexologist, he must deny the significance of drives, but he is not able consistently to maintain such a position.

3. Although severely critical of Freudian psychoanalysis, the author, in making impulse-inhibition and moral over-restraint the cause of neurosis, outfreuds Freud.

4. Although it is not entirely fair to characterize "conditioned reflex therapy" as symptom-treatment, there is much to justify such a characterization. Symptoms are not seen as solutions (habits) which the patient has developed as means of dealing with anxiety, but as "inhibitory conditioned reflexes." By means of "auto-hypnosis" (shades of Coué) which is based on (unreproducible) experimental findings reported by Huggins, the patient is given "exercises" which are supposed to replace the inhibitory reflexes by excitatory ones.

5. Sounder but less stressed parts of Mr. Salter's therapeutic procedure include: (a) encouraging the patient to engage in "feeling-talk"—which is reminiscent of James' admonition never to have an emotion without acting upon it; (b) emphasis upon changing self-attitudes by altering one's relations with other significant persons (cf. Sullivan); and (c) advice on breaking "vicious circles"—but the author here fails to face some of the implications of his own argument.

Conditioned Reflex Therapy is a defiant, impudent book! But who dares admit to complacency in respect either to theory or practice in the field of psychotherapy today? The author's approach is directive to a degree which will startle analysts only a little less than it will client-centered therapists: "I then become stern. . . . I am the authority. I will decide" (p. 69). "You'll do just as I say. Do you understand?" (p. 169). "I'm running the show" (p. 181). But unless one dismisses as fraudulent the case histories and the changes in Bernreuter scores which the author reports, one must draw the inference that Mr. Salter is achieving results which are neither much better nor much worse than those of more conventional therapists.

Here, again, we face that pertinacious question: What *is* psychotherapy—its processes and principles? By the art of simplicity combined with bristling self-assuredness, the author makes out a case which the unwary will mistake for an answer.

O. H. MOWRER.

University of Illinois.

DERI, SUSAN K. *The Szondi Test.* New York: Grune & Stratton, 1949. $5.00.

The publication of Susan Deri's long awaited book on the Szondi method will, in the opinion of this reviewer, be to the Szondi what Beck's *Manual* was to the Rorschach in 1937. Both books were written by ardent disciples of a master; both bear the stamp of their author's original extension of a creative idea; and both were produced in response to the intense interest of a few at a stage of apparent readiness for sharing with the many.

Since Deri began her seminars in 1945, research under her leadership has grown in an atmosphere of mingled fascination and skepticism. American psychologists have found it impossible to take seriously Szondi's theory of choice based on affinity between the genes of a subject and the genes of a patient whose photograph is selected as "liked" or "disliked." The test has nevertheless had strong appeal, not only because of its ease and simplicity of administration and recording, but because the experience of many clinicians suggests that clinically valid interpretations of choice profiles can be derived even without fully acceptable explanation of the assumptions developed from this theory.

In her book, Deri loyally states Szondi's views, including the typology which appears in his own untranslated book. Then, with a tact which is unusual in professional writing, she sets his theory aside and proceeds to the description of her own rationale of interpretation. To this task she brings her experience as a practicing analyst, and her training under Lewin. The determinants of choice are discussed not in terms of genetic characteristics of pictured patients falling in eight nosological categories, but rather in terms of stereotypes which the pictures represent to subjects who select or reject them in terms of their own needs. Choice, according to Deri, is the result of tension or satiation in eight need-tension systems, grouped into four vectors, which provide "an octagonal gauge for the understanding of personality dynamics." Although the need systems are reconciled both with Freudian constructs and with topological conceptualization of personality structure, the vectorial approach offers a new set of dimensions within which to conceive the dynamic processes of sexuality and aggression, cycles of control and discharge, ego structure and function, and object relationship.

To the extent that the clinical validity of the test proves verifiable, the Szondi test, as Deri describes it, promises several unique contributions. Through serial administration, it is possible to study test-retest change, thus capitalizing on a form of unreliability which in more cumbersome methods is an obstacle. Second, the interpretation of "loaded" and "open" factors, in terms of discharge and discharge readiness, allows the prediction of symptoms and overt behavior patterns in relation to stable, "root" factors in which may be sought the deeper, inner springs of action.

Although Deri's book does not in itself clear the mystery of why the test appears to "work," it offers a clear formulation of assumptions which should lend themselves to experiment. It is to be hoped that clinicians will not use the book as a manual of signs for the deceptively easy identification of syndromes and traits, but will utilize it as a source book of hypotheses for systematic exploration.

HELEN D. SARGENT.

Topeka, Kansas.

KORNER, ANNELIESE FRIEDSAM. *Some aspects of hostility in young children.* New York: Grune & Stratton, 1949. Pp. ix+194. $3.50.

This is a report of a study of hostile behavior in 20 nursery school children (11 boys and 9 girls) between the ages of 4 and 6 years. Its main value lies in the author's use of several different methods of obtaining the data, and her careful comparison and evaluation of behaviors and ratings of behaviors obtained under different conditions and by different methods.

All of the 4 to 6 year old children in one nursery school who were living at home with both parents were used in the study. They were predominantly Roman Catholic, of below average socio-economic status. In the first experimental situation each child was shown furniture and doll families representing his own family constellation. This material was used in telling a series of ten stories. The stories represented situations involving conflicts of the type which are prevalent in families with small children. Each story conflict was left unresolved, with the child requested to complete it, both in words and actions with the toys. In a later play session these and additional toys, as well as clay and crayons, were made available for the child to play with as he chose. Interviews with parents included questions about the children's actual behavior in situations like those described in the stories. The children were also rated by their nursery school teachers on their hostile behavior in school, and by the experimenter from her observations during the experimental play periods.

No correlation was found between the children's displays of hostility observed or rated in the different situations. The children could be classed, roughly, into four groups: those who were consistently hostile, those who consistently showed little hostility, those who were hostile in real life but mild in the play situations, and those who enacted very hostile play behavior but were mild in real life situations. Between or within these groups there was no clear tendency for differences in the adequacy of their emotional adjustment, for acceptance or rejection by their parents, or the parents' method or severity of punishment.

However, analysis of all the case material on a given child proved revealing of the dynamics of that child's behavior. There was evidence that those children who were rejected were profoundly affected, although their reactions varied.

The study is well done and is in general excellent. In only one instance does the author appear to make generalizations not warranted by her material. She assumes (p. 141) that these children have an unusually high number of symptoms of emotional maladjustment, and attributes this high incidence to their age. However, she gives no supporting evidence for either assumption. She cites no comparative data on incidence for other children, other ages, or children of different cultural or economic backgrounds. This particular point is a minor one, however, and does not affect the main conclusion of her study. A practical application of this conclusion is a warning to clinicians that a particular kind of behavior, such as hostility, should be interpreted with caution, and only in the light of other relevant information.

University of California, Berkeley. NANCY BAYLEY.

BROUWER, PAUL J. *Student personnel services in general education.* Washington, D. C.: American Council on Education, 1949. Pp. xix+317. $3.50.

This volume is a report of some experiences in developing personnel programs in twenty-two colleges associated together in a cooperative study in general education. The point of view is stated on the book cover as: "all those who influence the educational experiences of students are personnel workers." On p. xi of the preface, a further identifying statement is the following: "analyzing ways to identify and satisfy the needs of students through the coordinated use of fact-finding devices of personnel services." A good deal of attention is given to a contrast between two types of counseling: (1) prescriptive emphases, identified with intellectual understanding by the counselor and communicated to the client; and (2) permissive emphases derived from the non-directive point of view and in which the atmosphere of the counseling interview facilitates growth of the individual client. The latter point of view has been adopted by the author.

Included in the volume are chapters on the counseling process, the counseling program, extracurricular life, living arrangements, pre- and post-college personnel services, specialized personnel services, personnel services in the classroom, and administration of personnel services. Two personnel or counseling instruments are described, a self inventory of personal-social relations and an inventory of counseling relations. Chapter 14 contains an outline of a personnel philosophy of education which, to this reviewer, is largely the expression of an individualization movement in education, though, to be sure, this present statement includes mental hygiene and other emotional growth aspects of instructions. Part III is entitled, "The Principles of Personnel Services" and includes a number of sections. Sister Annette writes on psychological principles; Professor Mahoney discusses physiological principles; a chapter on philosophical principles is contributed by Professor Cannom; and a chapter on the sociological principles of personnel work presumably prepared by the author, Mr. Brouwer. These four interesting chapters make more explicit some assumptions regarding the nature of human nature pertinent to personnel work. These assumptions are identified in this volume: the unity of the individual; individual differences; and individual motivations as goals. In the chapter on sociological principles, socialization is described in terms of interiorizing the social environment from childhood to adulthood. The campus is spoken of in general terms as sub-cultures. Sister Annette defines adaptive counseling, ". . . which aims to help the student to help himself. It is characterized by an adaptation of procedures to meet the particular needs of the student as they are revealed in the counseling interview." This is presented as a formulation mediating between a prescriptive and permissive emphasis in counseling theory.

This book is an interesting contribution to the meager literature of explicit emphases, points of view, and consequent experiences in the development of institutional programs of personnel services to students.

E. G. WILLIAMSON.

University of Minnesota.

VERNON, P. E., & PARRY, J. B. *Personnel selection in the British Forces.* London: Univ. of London Press, 1949. Pp. 324. 20 Shillings net.

The authors have linked up the developments of the last war in selection and guidance in the British Forces with the broader aspects of vocational and educational psychology. Some 80 pages (Part I) are concerned with the organization of selection procedures of the Navy, Army, Women's Corps and Air Force. Then the scientific contributions of this work are included in general discussion of familiar topics of personnel psychology. This forms the major portion of the book (Part II).

The book is written primarily for the Britisher, to show the industrialist and educationalist the usefulness achieved by psychology during the war years. It is limited in scope pretty much to British sources and it was not intended by the authors to survey American war-time psychological work.

The reader of this book will secure a general picture of work accomplished in personnel selection in the various British Forces. He will not find inclusive statements of war-time research projects. But their results are extensively quoted in connection with various topics in vocational psychology. The main RAF findings on selection are summarized in Ch. XVI. A useful reference will be found in Appendix II which lists by title and author, "The Main Psychological Tests Used in the Forces," giving the number of items, time of administration and reliability of tests.

The authors maintain the point of view that the "clinical" as contrasted with the "psychometric" approach yields a truer account of the candidate. There is an extensive discussion of the use of the interview (Ch. IX) and it is stated that "The interview rather than tests must remain the prime instrument of vocational classification" (p. 164).

Various "problems" indicated in the text will strike a chord of sympathy in American war-time researchers as in matters of military administration, "Without belittling in any way the fine work by all types and grades of staff, it must be admitted that the organization did not always function smoothly. Psychologists suffered considerable frustration. They were commonly of lower rank. . . . Policies they advocated as scientifically sound were often rejected, and the methods they devised were often misapplied and misinterpreted . . . " (p. 42); or in technical matters, "Some training courses were largely inappropriate; training devices and visual aids were employed without proof of their value; old-fashioned examination methods were used which not only probably failed the wrong men, but also provided thoroughly unsatisfactory criteria against which to gauge the validity of selection procedures" (p. 98).

The authors include an excellent summary of conclusions based on personnel selection in the Forces which they regard as having vocational and educational applications in peace-time. No new psychological methods or techniques of selection are described, but added factual material is supplied for their understanding.

The book is packed with facts from the work of selection in the Armed Forces of Great Britain and it will probably be the only available general source for this material. An extensive index has been prepared for reference work. By intent, the book does not deal with the work of psychologists on problems of training, the design and layout of equipment and the study of morale.

New York University. DOUGLAS H. FRYER.

Mursell, James L. *Psychological testing.* (2nd Ed.) New York: Longmans, Green, 1949. Pp. xvi+488. $4.00.

A large part of this book (198 pages) is devoted to descriptions and critical comments on some 100 well-known tests of intelligence, aptitude, personality, interest, attitude, and character. This central core of the book is preceded by discussions of the general characteristics of mental tests and certain basic considerations affecting their use, such as validity and reliability, as well as by a chapter on the concept of general intelligence. It is followed by a treatment of mental growth and its relationship to heredity and environment, socio-economic factors, family relationships, schooling, and race. Constancy of the IQ, the distribution of mental traits, and the psychological significance of test scores are also discussed. The book ends with a chapter on the evolution and improvement of mental testing.

The author writes well and has read a good deal of the literature in the field of measurement; the text is liberally sprinkled with footnote references to such sources. It is probable that students will find the book readable and informative.

In the opinion of the present reviewer, however, the book is not scholarly in any sense of the term. It is easy enough to point out specific errors of fact and interpretation in it, but it is time consuming to demonstrate how fundamental misconceptions and misunderstandings of psychological testing have led to the subtle distortions and misrepresentations that are woven inextricably into the content. It reads as though an intelligent layman with a flair for writing but with no background in psychology or statistics had surveyed the literature on testing, taken a look at a hundred of the most commonly used tests, and then dashed off a book.

To illustrate the sorts of errors that characterize the text, let us consider the section on reliability in chapter 2. This section starts off "'If unchanging subjects are measured twice with a perfectly reliable instrument by a perfectly reliable agent, the correlation between the two sets of scores is 1.00' (Walker, p. 365). This statement briefly summarizes what is meant by reliability of measurement." The quotation from Walker is true, but it does not at all summarize what is meant by reliability of measurement. The classic definition of reliability by Spearman refers to the correlation of two separate but hypothetically equivalent measures. Later in the chapter (at the bottom of page 47), the writer contradicts an essential element in the formulation of reliability in terms of analysis-of-variance techniques when he writes, "Interdependent items, i.e., those which present the same problem in different forms tend to lower reliability." On the contrary, as item intercorrelation is increased by testing the same point in several ways, test reliability tends to increase.

On page 54, the author computes a standard error of measurement, finds it to be 9.36, and then interprets it as follows: "This tells us that there is a 68% chance that a person making any given score on the first testing will score within a range of +9.36 to −9.36 of that score on the second testing. That is, if a person makes a score of 170 on one testing, there is a 68% chance that his score on another testing (which is often called his 'true score,' i.e., his score on *any other* testing) will fall between 161 and 179." These statements will not meet with the approval of educational statisticians.

If anyone supposes that the interpretation of test scores for practical purposes can safely be made by the author in spite of his apparent misunderstanding of the concepts of reliability, he should examine the statements at the bottom of p. 110 and the top of p. 111 concerning the interpretations of IQ's derived from the revised Stanford-Binet Scales. The author states, "Thus general reliability coefficients cannot be worked out for this scale, or for any such scales." The truth is that a general reliability coefficient for Stanford-Binet IQ's could easily be computed, but standard errors of measurement at several IQ levels are much more valuable for the informed test user than the general reliability coefficient.

In the reviewer's opinion, the faults of this book greatly outweigh any merits that it may possess. It does not seem to come up to the standard of the author's previous writings or to accomplish its stated purpose; in the reviewer's judgment, it should never have been published.

Hunter College. FREDERICK B. DAVIS.

BOOKS AND MATERIALS RECEIVED

ALLPORT, GORDON W. *The individual and his religion.* New York: Macmillan, 1950. Pp. xi+147. $2.50.

BALDWIN, ALFRED L., KALHORN, JOAN, AND BREESE, HUFFMAN. *The appraisal of parent behavior. Psychological Monographs* No. 299. (Vol. 63, No. 4.) Washington, D. C.: American Psychological Assn., 1949. Pp. vii+85. $1.50.

BOGERT, CORNELIA H. *With brushes of comet's hair.* New York: Exposition Press, 1950. Pp. 165. $5.00.

BOND, GUY L., AND WAGNER, EVA BOND. *Teaching the child to read.* New York: Macmillan, 1950. Pp. xi+467. $3.75.

BRITT, STEUART HENDERSON (Ed.). *Selected readings in social psychology.* New York: Rinehart, 1950. Pp. xvi+507. $2.00.

CHEVIGNY, HECTOR, AND BRAVERMAN, SYDELL. *The adjustment of the blind.* New Haven: Yale Univ. Press, 1950. Pp. xvi+320. $4.00.

COCHRAN, WILLIAM G., AND COX, GERTRUDE M. *Experimental designs.* New York: Wiley, 1950. Pp. ix+454. $5.75.

Doppelt, Jerome Edward. *The organization of mental abilities in the age range of 13 to 17*. New York: Bureau of Publications, Teachers Coll., Columbia Univ., 1950. Pp. x+86. $2.10.

Dorcus, Roy M., and Jones, Margaret H. *Handbook of employee selection*. New York: McGraw-Hill, 1950. Pp. xv+349. $4.50.

Earnest, Ernest. *S. Weir Mitchell: novelist and physician*. Philadelphia: Univ. Pennsylvania Press, 1950. Pp. vii+279. $3.50.

Freeman, G. L., and Taylor, E. K. *How to pick leaders*. New York: Funk & Wagnalls, 1950. Pp. vii+226. $3.50.

Garrett, Henry E. *Psychology*. New York: American Book Co., 1950. Pp. ix+323. $3.00.

Garrison, Karl C. *The psychology of exceptional children*. (Rev. Ed.) New York: Ronald Press, 1950. Pp. xvii+517. $4.50.

Guilford, J. P. (Ed.) *Fields of psychology*. (2nd. Ed.) New York: Van Nostrand, 1950. Pp. xiv+779. $5.00.

Guilford, J. P. *Fundamental statistics in psychology and education*. New York: McGraw-Hill, 1950. Pp. xiii+633. $5.00.

Gurvitch, Georges. *Sociometry in France and the United States*. New York: Beacon House, 1950. Pp. vii+261. $7.50.

Guthrie, Edwin R., and Powers, Francis F. *Educational psychology*. New York: Ronald Press, 1950. Pp. vi+$4.00.

Hall, Victor E. (Ed.) *Annual review of physiology*. Vol. XII. Stanford, Calif.: Annual Reviews, Inc., 1950. Pp. vii+689. $6.00.

Halliday, James L. *Mr. Carlyle, my patient*. New York: Grune & Stratton, 1950. Pp. xiii+227. $3.50.

Hemphill, John K. *Situational factors in leadership*. Bureau of Educational Research Monograph No. 32. Columbus: Ohio State Univ., 1949. Pp. xii+136. $2.50. (Cloth, $3.00.)

Hurlock, Elizabeth. *Child development*. (Second Ed.) New York: McGraw-Hill, 1950. Pp. xvi+669. $4.50.

James, William. *The principles of psychology*. (Volumes I and II bound together.) New York: Dover Publications, 1950. Pp. 688. $7.50.

Kent, Grace H. *Mental tests in clinics for children*. New York: D. Van Nostrand, 1950. Pp. xii+180. $2.45.

Kirkpatrick, Clifford. *Religion and humanitarianism*. *Psychological Monographs* No. 304. (Vol. 63, No. 9.) Washington, D. C.: American Psychological Association, 1949. Pp. v+23. $.75.

Lawton, George. *How to be happy though young*. New York: Vanguard Press, 1949. Pp. xx+300. $3.00.

Lundin, Robert W. *The development and validation of a set of musical ability tests*. *Psychological Monographs* No. 350. (Vol. 63, No. 10.)

Washington, D. C.: American Psychological Assn., 1949. Pp. iii+20. $1.00.

MAY, ROLLO. *The meaning of anxiety.* New York: Ronald Press, 1950. Pp. xv+376. $4.50.

MICHAEL, WILLIAM B. *Factor analysis of tests and criteria. Psychological Monographs* No. 298. (Vol. 63, No. 3.) Washington, D. C.: American Psychological Assn., 1949. Pp. v+55. $1.00.

NEWCOMB, THEODORE M. *Social psychology.* New York: Dryden Press, 1950. Pp. xi+690. $4.50.

PLANT, JAMES S. *The envelope.* New York: The Commonwealth Fund, 1950. Pp. 299. $3.00.

POLLAK, OTTO. *The criminality of women.* Philadelphia: Universit of Pennsylvania Press, 1950. Pp. xxi+180. $3.50.

PROTHRO, E. TERRY, AND TESKA, P. T. *Psychology—a biosoci study of behavior.* New York: Ginn & Co., 1950. Pp. ix+546. $3.75.

SCHAER, HANS. *Religion and the cure of souls in Jung's psycholog* New York: Pantheon Books, Inc., 1950. Pp. 221. $3.50.

SPEARMAN, C., AND JONES, Ll. WYNN. *Human ability.* Londo Macmillan & Co., Ltd., 1950. Pp. vii+198. $2.50.

STONE, C. P. (Editor) *Annual review of psychology.* Stanfor Annual Reviews, Inc., 1950. Pp. ix+330. $6.00.

TAFT, DONALD R. *Criminology.* (Rev. Ed.) New York: Ma millan, 1950. Pp. xiv+704. $5.50.

WARNER, SAMUEL J. *The color preferences of psychiatric grou Psychological Monographs* No. 301. (Vol. 63, No. 6.) Washington, D. C American Psychological Assn., 1949. Pp. v+25. $0.75.

ZAHL, PAUL A. (Editor) *Blindness.* Princeton: Princeton Unive sity Press, 1950. Pp. xvi+576. $7.50.

ERRATUM

In listing the title of Muriel Potter's monograph in the Janua 1950 issue of THIS JOURNAL, The word "reception" was erroneousl used for "perception." The correct citation is as follows:

POTTER, MURIEL CATHERINE. *Perception of symbol orientation an early reading success.* Teach. Coll. Contr. Educ. No. 939. New Yor Bureau of Publications, Teach. Coll., Columbia Univ.. 1949. viii+69. $2.10.

Vol. 47, No. 5 September, 1950

Psychological Bulletin

STUDIES IN THE PSYCHOLOGY OF CHILDREN'S DRAWINGS: II 1928–1949

FLORENCE L. GOODENOUGH AND DALE B. HARRIS
Institute of Child Welfare, University of Minnesota

GENERAL REVIEWS AND SUMMARIES

More than twenty years have elapsed since the publication in this JOURNAL of a review by Goodenough (**105**) of the earlier literature on children's drawings. Although other writers have dealt with special phases of this subject since that date and Graewe's (**109**) *Geschichtlicher Überblick* covers the literature from 1903 to 1935 fairly completely, no recent complete summary, written in English, has come to our attention. Perhaps the nearest approach is the chapter by Meier (**199**) in the *Fortieth Yearbook of the National Society for the Study of Education* which includes a bibliography of 150 titles, the greater number of which deal with the art products of children. Goodenough's chapter on children's drawings (**106**) which appeared in the first edition of Murchison's *Handbook of Child Psychology* presents a more detailed discussion of the topic than that given in the earlier article, but this book is now out of print. Less complete reviews and summaries have appeared in connection with many of the books and articles mentioned in later sections of this article.

Of the many discussions of special aspects of the topic, by far the most elaborate and detailed is the series of four articles on the art of the abnormal published by Anastasi and Foley (**11**, **12**, **13**, **14**). Although the material deals mainly with adults, some data on the art of neurotic and psychotic children and adolescents are included. The chapter on drawing and painting in Bell's *Projective Techniques* (**25**) includes table, 25 pages in length, in which the various aspects of drawing ted in one or another of the studies included in his bibliography of titles are summarized together with their presumed clinical interetation. Bell cautions that drawings must be interpreted as wholes rather than in terms of separate aspects or parts, the meanings of which may change with the circumstances. Probably his table is most useful as a summary of hypotheses to be tested further, rather than as a diagnostic or clinical aid. The table includes appropriate references to the bibliography in the case of all entries, but it gives no indication of the comparative dependability of the items. Findings of carefully designed quantitative researches appear alongside shrewd (and sometimes naive) hunches based on single cases. Another summary of projective methods, including painting and finger painting, appears in the monograph by Lerner and others (**162**).

Bell's chapter provides a striking illustration of the rising tide of interest in drawing as a "projective method." The belief that the art of children is primarily a language, a form of expression, is by no means new, but its emphasis has shifted. Earlier studies were for the most (I) part based on the oft-repeated statement: "A child draws what he knows, rather than what he sees." At the present time many students of child art would revise this hypothesis to read: "A child draws what (II) he feels, rather than what he sees or knows to be true." In spite of the charlatans who have exploited the method and others with poor scientific training who have used it unwisely, evidence that the child in his drawings frequently gives outward expressions to his inner life of thoughts and feelings, to his fears and his desires, to his hopes and his frustrations, is steadily accumulating.

Although projective theory has provided the keynote for many of the recent studies in this field, more conventional topics have continued to attract the attention of many persons, especially during the first half of the period covered in this review. Studies of the developmental stages in drawing have confirmed earlier findings and have added many details to the general outline. The use of children's drawings as aids to the solution of various psychological problems, particularly those having to do with visual perception and with memory for what is seen, has become increasingly popular. Many of these reports include fairly comprehensive bibliographies. The use of drawing as a means of studying the divergences of various cultural or racial groups is at least a century old, but the recent studies are more sophisticated both in methodology and in the conclusions drawn. An excellent survey of methods with particular reference to those used in the study of the perception of form has been given by Cain (**50**). The literature on special artistic talent was reviewed by Meier (**194**) a number of years ago and in connection with a later series of studies carried on under his direction at the State University of Iowa (**195, 198**). Other reports have considered such questions as art eduction in the schools, the measurement of ability in art, the extent of art appreciation, and the relation of talent in art to other abilities or personality traits.

The brief accounts of the literature which follow are based upon reading of the original book or article by one or both of the present writers when the sources were available to them. When the originals were inaccessible, the reports in the *Psychological Abstracts* have been used. In all such cases a reference to the abstract is appended to the citation in the Bibliography.

THE DEVELOPMENT OF GRAPHIC EXPRESSION IN CHILDHOOD AND ADOLESCENCE

General Developmental Trends. The developmental aspects of child art will be considered first, not only because interest in this topic preceded interest in most of the others in point of time, but for the more important reason that a knowledge of the usual is essential for the recognition of the unusual. Many of the reputedly abnormal features in the drawings of single children or of small, selected groups lose their apparent significance when the age and sex of the subjects and the conditions under which the drawings were made become known.

Both cross-sectional studies of the drawings of large groups of children at succeeding ages and longitudinal studies of the work of a single child over a period of years have been employed to examine typical age changes in the form and content of graphic expression. Eng's very complete and detailed study of the drawings of her little niece, originally published in German, is now available in English translation (**83**). Hildreth's study (**127**) of more than four thousand drawings made by a boy between the ages of two and eleven years adds a third to the list of intensive studies of the drawings of individual children that have appeared in the literature. McCurdy's study (**181**) of more than one thousand drawings made by a boy between the ages of two and one-half and seven years will be referred to later in this paper since it relates only to the Goodenough technique of scoring the drawings of a man to obtain an estimate of intelligence.

Hildreth's monograph is of interest for several reasons. In the first place the drawings were made by a boy whereas the children studied by Luquet and Eng were girls.[1] More than half of the drawings made by the boy were attempts to depict trains; those made by the girls in the two earlier series were chiefly pictures of human beings, houses, animals, trees, and flowers. The difference in content renders it possible to examine the three sets of data in terms of their psychological aspects rather than in terms of their particular details. In analyzing the train drawings made by her subject, Hildreth lays particular stress on the "wholeness" of the concept which the child was attempting to depict. At no time was there evidence of a building up of the whole from isolated parts; always the *train* was primary. As knowledge of the various parts and their functions increased and became more specific, these parts were indicated with greater accuracy of form, position, and number. A growth curve is presented, showing the accumulation of new features added each year. As has been found true of most developmental traits, improvement is most rapid during the early years, with negative acceleration becoming clearly apparent after the age of five. Although Hildreth emphasizes the evolutionary character of the changes that occur with advancing age, she nevertheless outlines sixteen stages in the development of the train drawing, each of which is marked by the appearance of some characteristic features not previously noted. In spite of the differences in subject matter, these changes are quite in line with those previously noted by Rouma, Luquet, Eng, and others.

---

[1] For a brief account of these studies see References **83** and **176** of the Bibliography at the end of this article.

The hypothesis that, in drawing, children proceed from the whole to the parts is not entirely confirmed by Barnhart (**19**), who constructed an ingenious device for studying this problem. It consisted of a frame in which the drawing surface covered a sheet of carbon paper which, in turn, covered a long sheet of paper wound on a roller. At intervals of one minute, while drawing, the subject was asked to pause while the lower paper was pulled out to the length of one panel. This made it possible to determine the order of drawing. Fifty-two children between the ages of five and sixteen years were asked to make a picture of children playing in the snow in a park. Some children, chiefly the older ones, began with a broad outline of the entire picture and later filled in the details; but the younger children more frequently completed each detail separately before passing on to the next. The complexity of the subject is likely to have been a contributing factor in this case but it would be interesting to know what characteristics other than age differentiated the two groups. Barnhart also examined the manner in which space was depicted by children of different ages. His findings are very similar to those obtained by Stern and others many years ago. At first there is no apparent concept of graphic portrayal of space. Objects are scattered at random over the paper.[2] This is followed by a linear form of representation in which objects are arranged in a row, usually upon a ground line. Next comes a period of ranked space in which there are two or more such rows with the ground lines in the background either straight or curved to indicate relative distance, but showing only rudimentary ideas of perspective. A transitional period follows leading to the final idea of "true space" with recognition of such factors as perspective, partial concealment of objects in the background by those in the foreground, foreshortening, and the like, although lack of skill may lead to very crude ways of depicting these factors.

Variations in style resulting from differences in the manner of apprehending part-whole relationships at different ages were studied by

---

[2] Even at this stage, association between persons or other objects is shown by placing them near each other or even in juxtaposition, but the concept involved seems to be that of a functional relationship, rather than that of position in space.

Kobayashi (**145**) and by Martin (**192**). Both found that with advancing age there is an increasing tendency to perceive an object or figure as a differentiated whole, rather than as a conglomeration of parts. Much the same results are reported by Katô (**139**) for Japanese kindergarten children who were asked to draw from a doll-model of a man placed before them. At first only the main parts were drawn, then details were added, until finally the concept of the figure as a whole with differentiated parts was indicated.

The prerepresentative stages of drawing or scribbling have received more attention during the past twenty years than formerly. Seeman (**271**), Schilder (**265**), and Bender (**26, 28**) stress the importance of the early circular movement—the "whorl"—which Seeman finds to be equally characteristic of the drawings of all infants or young children, regardless of race or culture. The whorl gradually evolves into the "stick-man" in which the circle still predominates but to which various appendages are added in accordance with cultural influences. The apparently universal character of the "whorl" leads Read (**254**) to a rather curious theory of the development of child art which appears to be a mixture of the descriptive accounts of the early writers on the subject with psychoanalytic theories and typology, particularly that of Jung. He advances the theory of "primordial images" exemplified by the circle, the whorl, and radiating lines appearing in the drawings of young children, and believes that these forms represent "electrostatic patterns produced in the cortex by normal phenomenal experience" (p. 189). To the present writers, this does not appear to add much to the generally accepted belief that not merely drawing but behavior of all kinds is mediated through cortical action.

To the very young child, "writing" and "drawing" are much the same. Legrün (**159**) describes a series of developmental stages by which the two forms of graphic expression become differentiated. The progress is from completely unorganized scribbles with no apparent design to forms of a more characteristic shape with occasional attempts at letters of the alphabet. Hildreth (**126**) also reports the early stages of handwriting shown by 170 children between the ages of three and six years. She is able to differentiate seven levels of progress beyond which further improvement seems to involve mainly a higher degree of accuracy, neatness, and speed of writing. Both Legrün and Hildreth note that letter-like forms usually appear earlier than true letters and that irregular scribblings tend to give place to arrangement in forms of lines or rows as the concept of writing is evolved.

The X-ray effect in drawing and confusion of full face and profile, which have been noted as immature or transitional stages in achieving representative drawing, receive a different interpretation from Lassen (**158**). He assigns these effects to the expression of a creative urge which expresses itself, with age, in a freer organization of the objective relation of things. Inasmuch as both of these characteristics are found more frequently in the work of feebleminded than of normals of the same mental age, the present writers are reluctant to accept his term "creative" in its ordinary meaning.

Among other investigations in which the findings have resulted mainly in a reiteration of developmental trends previously observed, may be mentioned the work of Beach and Bressler (**22**), Hetzer (**124**), and Zambaldi (**326**), who have emphasized that developmental tendencies be considered when planning a program of art education for children. A similar viewpoint has been expressed in many brief articles and reviews in encyclopedias, educational journals, and the like not mentioned here.

Developmental Direction and Orientation in Drawing. The development of directionality in drawing was the subject of an extensive investigation by Gesell and Ames (**102**). Approximately fifteen hundred drawings were secured from children between the ages of eighteen months and seven years, most of whom were tested semi-annually, and from twelve adults. The tasks included drawing a vertical and a horizontal stroke in imitation of the examiner and drawing a cross, a circle, a square, a rectangle with a diagonal, and a diamond from copy. Of the reported findings, the following seem most reliably established: (a) For both the vertical and the horizontal strokes the average length of line increases as age advances. This probably reflects more purposeful action as well as improved control in use of the hand. (b) Symmetry in the form of the cross improves with age as does the tendency to make the cross by means of two intersecting lines instead of drawing one or both half-lines to the point of intersection and completing the figure by means of a separate line. (c) Younger children commonly draw both the square and the diamond by means of four separate lines. Older children and adults usually draw the square by means of a single continuous line but are likely to use four downward strokes for the diamond. The article also notes the usual errors in copying the square and the diamond, but these are too well known to require description here. No consistent relationship between directionality and handedness is indicated. Rice (**256**), however, who had 293 subjects ranging in age from two and one-half to fourteen years copy the diamond both in the horizontal and the vertical position, found that right-handed subjects commonly follow a clockwise system in making the strokes. Although her number of left-handed subjects was too small to yield highly dependable results, the indicated trend is in the opposite direction. Other findings are similar to those obtained by Gesell and Ames. Hicks and Ralph (**125**) found no reliable difference between the final performance of an experimental group of nursery school children who were required to trace the Porteus Diamond Maze twenty times with the preferred hand and that of a control group who traced it only six times. As a matter of fact, the latter did slightly better than the former; a fact which may be attributed either to chance or to loss of interest on the part of the experimental group.

Orientation in the direction depicted in the drawings of railway trains by kindergarten and primary grade children in Japan and Formosa was studied by Iinuma and Watanabe (**136**) who present the following figures: Out of a total of 1,648 drawings, 26 were front views and the direction of 105 could not be determined. Of the remainder, more than 80 per cent were shown as running toward the left. Hildreth (**127**) reports a similar tendency in the drawings of the boy whom she studied but does not give exact figures. Unwillingness to cover their work with the hand that holds the pencil on the part of children of whom the majority are right-handed, seems the best explanation for this tendency. This point has been particularly stressed by Østlyngen (**236**). A similar trend has been noted in drawings of animals and in profile drawings of human figures. Both usually face toward the left. Zesbaugh (**327**) notes this in her analysis of more than ten thousand drawings of a "mail-man" made by elementary school children. She notes also that purely schematic drawings have pretty well disappeared by the time the children reach the fourth school grade and that some attempts at depicting the special characteristics of the postman's uniform appear as early as the second grade. Further developmental trends similar to those noted by other investigators are also reported.

Graewe's very intensive study (**107**) of over six thousand drawings by subjects ranging in age from the preschool to the adult level included both free drawings and drawing from a model with later reproduction from memory. His treatment of the data is descriptive only; there are no tables and few quantitative statements of any kind. However, his long and intensive study of the problem enables him to make many keen observations with respect to developmental changes relating to such matters as the tendency to perseveration, the trend toward concreteness, the manner of representing space or perspective, foreshortening, and the like. In a second study (**108**) Graewe examined the developmental changes in the child's manner of depicting animals. At first there is no distinction between the animal drawing and the drawing of the human figure; the child differentiates them only by naming. Later on a change in orientation appears, with the animal's body usually placed in the horizontal position, the human figure in the vertical position. Still later some distinguishing appendage is added, such as a tail or a larger number of legs which in time is reduced to four. In the final stage, emphasis is placed on the total outline; the animal is an animal from the start. Kerr (**140**), who studied children's drawings of houses, found few consistent qualitative changes with age, but a general improvement in symmetry and in the neatness of lines and angles. The absence of marked qualitative differences is probably attributable to the wide variance in the type of houses with which children are familiar. On the other hand, Neubauer (**231**) reports that adolescents, when asked to decorate and complete an outline drawing of a teapot, show a distinct age progression in use of color, geometric patterns, composition, and the emergence of idealistic and realistic types. These findings differ markedly from those generally reported for adolescents who show no particular aptitude for graphic art.

Cases of inversions in the orientation of children's drawings are reported by Belart (**24**), Billing (**35**), Gerald (**101**), and Pearson (**240**). For the most part, interpretations follow those generally given for inversions in the writing of young children. Although Pearson regards these cases as being extremely atypical at the ages specified (five and six years), they are actually less uncommon than he supposed. Many children, like some of the more primitive groups described in another section of this paper, have difficulty at first in identifying the idea of "up" with the top of the paper.

Child Art: Drawing from a Model. A number of persons have been interested in the character of the changes made when a child attempts to reproduce a design from copy, either directly or from memory. Graewe's (**107**) study has already been mentioned. By far the most elaborately analyzed of these studies was made by Portocarrero de Linares (**248**), whose subjects were 1,375 girls from the public schools of Lima, Peru, ranging from kindergarten through "preparatory" to Grade VI, eight levels in all. Each subject was asked to copy a series of eight designs which were printed on the test blank. Most of the designs were highly conventionalized representations of common objects such as a table or a fish. The results are presented by grade levels rather than by ages. The following eight categories are considered: (1) Entirely correct reproduction, (2) correct except for minor errors in proportion or symmetry, (3) amplification, i.e., the addition of elements not present in the model, (4) simplification, (5) deformation, (6) broken structures in which elements that should have been connected are drawn separately, (7) concretization, in which changes were made in the conventionalized figure in order to bring about a closer resemblance to the real object, and (8) confabulation in which the subject apparently drew pictures at will, with no attempt to copy the model. As might be expected, a general improvement with advance in school grade is noted for all the models, but a comparison of changes in the various categories is interesting. In all grades and for all models, simplification is much more common than amplification. In Grade VI, for example, the relative proportions are 12.9 per cent to 0.6 per cent. Concretization and confabulation are largely confined to the kindergarten and preparatory classes except in the case of the much-conventionalized fish, for which the percentages in the former category continue to make up an important percentage of the total throughout the first four grades. Portocarrero de Linares emphasizes as does Hildreth (**129**), who conducted a similar experiment with four-year-old New York City children, the tendency to simplify that which is too difficult or to give meaning to that which is meaningless. Sorge's (**279**) work with rural and urban European children and Homma's (**131**) with Japanese children lead to a similar conclusion and suggest that the tendency may be well-nigh universal. Paulsson (**239**), it may be remembered, found the same thing to be true for adults.

Burkhardt (**48**) describes an experiment in which a comparison was made between the attempts of children to copy drawings of familiar forms with their copies of meaningless figures having approximately the same number of lines and angles. Alterations of form, which in Burkhardt's opinion represent "improvements" from the child's point of view, were much more common in the first instance than in the second. Apparently, unless a child is able to see some meaning in a figure at the start, he is not likely to try to improve upon it. Although Portocarrero de Linares does not classify her figures with respect to their meaningfulness, it may be noted that her Models II and VIII, which to us seem to be the least abstract in her series, do tend to show more alterations in the direction of concreteness than the others.

It is apparent that no universal statement can be made with respect to the relative tendency to add or to omit elements in drawing, for this will vary with the difficulty and meaningfulness of the material. Moreover, difficulty and meaningfulness are not absolute factors, but depend on the level of development of the subjects. Under some conditions, simplification may be rare and amplification for the purpose of giving a more obvious meaning may be the rule. Burton and Tueller (**49**) found this to be the case when they asked children of nursery school age to copy a circle with three dots so arranged within it as to suggest a face. The subjects were urged to make successive copies until they refused to make more. The prototype of this "satiation" experiment is, of course, the work of Lewin (**166**), who compared the performances of normal and feebleminded children. Bender (**27**) conducted a similar experiment with children suffering from organic brain damage. All these investigators find that when it serves to give concrete meaning to the relatively abstract symbol, the tendency to amplify is more common than the tendency to simplify. Apparently, however, this is the case only when the original stimulus is well within the child's ability to copy, and seems to him an imperfect or incomplete representation of the object he conceives it to depict.

In spite of the complexity of the task of discovering under which conditions children amplify in copying designs and under which they simplify, more research should be done in this area. The work of Engerth (**84**), of Vedder (**302**), and of Bender (**27, 28**), suggests that brain damage in children as well as in adults may be indicated by distortions in drawing and copying. Mira's Myokinetic Psychodiagnosis

(**206**), although promising for work in this area, has apparently not been tried out with children. The Ellis visual designs, similar to Bender's figures, have been used with children by Wood and Shulman (**323**) and by Lord and Wood (**171**). From these studies it is clear that a great deal more must be learned about the distortions which are characteristically introduced by the immaturity factor in drawings of normal children of different ages before we can be warranted in drawing conclusions in respect to their meaning for the abnormal.

Rostohar (**258**) had elementary school children reproduce from memory a number of designs, some of which were in black and white, others in color. After each unsuccessful attempt the original was again exposed and a new trial given. This was continued until an approximately correct copy was made or until it became apparent that no further improvement was likely. Examination of the successive trials showed that two different methods were employed in building up the memory image. Some children proceeded from the general contour to the details while others worked from the parts to the whole. Retests at weekly intervals indicated that the process of forgetting followed in reverse order the same general plan as the process of learning. Children belonging to the first group lost the details before the general form was forgotten; those of the second group retained details after the main figure could no longer be recalled.

Two papers by Slochower (**277, 278**) describe the attempts of children between the ages of five and ten years to reproduce line drawings of geometrical figures in each of three media—pencil drawing, clay, and Tinker Toy. The younger children kept to the two dimensional forms shown in the models with only such minor variations as were necessitated by the medium used. Children of nine and ten, however, showed a definite tendency toward the construction of three-dimensional forms when the clay and Tinker Toys were used. Slochower also found, as have others, a definite tendency to simplify the more difficult forms and an increase, with age, in accuracy of reproduction. Nakae (**218**) posed the task of copying a solid model of a cylinder for four- and five-year-olds. The abstract does not indicate that any attempt to depict three dimensions appeared.

Lark-Horovitz (**154**), who compared children and adults with respect to their ability to draw a series of models from memory, found that although on the average the adults did better than the children, there were many who drew schematically as children do. She comments particularly on the "interlinkage" of different sensory modalities in producing the memory image because of which the drawings of different persons will vary in accordance with the particular type of sense impression that was dominant at the time the image was formed.

Child Art: Formal Elements. Ellsworth (**82**) utilized the free easel paintings of 20 nursery school children as a means of studying the early stages of compositional design, particularly with respect to the relative use of line and mass. She found that if design was apparent, it most frequently took the form of a simple element placed in rows, columns. concentric circles, or in attempts to reproduce figures. Lines were more often used for the figures; mass more often when filling in the background. A more extensive study in this area was made by Cockrell (**57**), who collected 1,550 paintings from three nursery schools, with a total enrolment of 60 children between the ages of twenty months and six years. Cockrell reports that the devices most frequently employed for graphic expression were contrast and opposition, which occurred in almost every painting; balance of line, mass, or color, which was observed in 637 paintings; repetition, which was noted in 262 instances; rhythm of movement found in 180 cases; and symmetry, found in 93 instances. All examples of the group last named were made by the same thirteen children. Unfortunately little information is given with respect to the abilities and personality traits differentiating these subjects from their classmates. Cockrell concludes that drill in design is unnecessary for young children.

Haubold (**121**) approached the question of children's appreciation of style by a more objective method than those frequently used. He utilized photographic copies of pictures by different masters. In each set of pictures the subject matter of all the paintings was similar but the artists differed. Children were shown two sets at a time and asked to match the pictures they thought had been made by the same person. Success improved steadily from 35 per cent at the age of eight years to 52 per cent at the age of sixteen. We have not seen the original of this article and the abstract, unfortunately, does not provide the information needed to determine how far these figures exceed chance. It would be worth while to repeat the study with American children.

Child Art: Foreshadowing "Projective Methods." Biber's study (**34**) of the drawings of children in the Bank Street School not only enabled her to verify many of the conclusions as to developmental trends reached by her predecessors, but also led her to draw certain conclusions with respect to the indications of personal-social characteristics shown in the drawings made by individual children. As is true of many such studies, however, the line between correlation and coincidence is not sharply drawn.

Luquet, whose many and valuable studies in the field of children's drawings are too well known to require much discussion here, after reaffirming and amplifying his earlier position in two articles published in 1927 and 1929 (**175, 176**) turned his attention to a problem that had long interested him—the relation of child art to primitive art (**177**). The two articles first cited are of interest chiefly because in them he places particular emphasis upon what he calls the stage of "intellectual realism," where his discussion definitely foreshadows the modern "projective methods." Lowenfeld (**174**) gives an account of the progressive stages in the development of drawing which in many respects is very similar to that outlined by Luquet, but like most of the modern writers he places even more stress upon the subjective aspects of the drawing than did Luquet. He points out that the drawings of the young child reflect his own desires, feelings, beliefs, and fancies rather than objective reality. Proportions, for example, represent values as the child conceives them rather than absolute size. Later on, as the child moves into what Luquet designated as the final period, that of realistic representation, his self-confidence is shaken (so Lowenfeld believes) by the contrast between his internal and his external worlds. He may then give up all forms of creative activity or he may express his conflicts and confusions by the character of his art products.

Finger painting has appealed to a number of persons as offering unusual opportunities for expressive work because of the direct and primitive character of the method of working with the material. A bibliography of work on this topic has been prepared by Napoli (**220**). Blum and Dragositz (**38**) attempted to study its developmental aspects by a comparison of the work of first-grade and sixth-grade children. Unfamiliarity of both groups with the medium may perhaps account for the fact that so few differences were found. The older children obtained somewhat neater results with less smearing of the colors at points of junction and better covering of the paper. Even these differences, however, were so small that two teachers of art who served as judges were unable to separate the work of the two groups with much better than chance success.

In her study previously mentioned, Cockrell noted that in practically all of the drawings made by the older children of her group, evidence of an idea or purpose was apparent. Hatano (**120**) found that young children were more free in their attempts at graphic representation of a story told to them than in their attempts to reproduce it verbally. Scheer (**264**) tabulated the subjects voluntarily drawn by adolescent boys in an attempt to ascertain the dominant interests characteristic of this period of development. Although his method of classification differs from the methods employed by earlier writers, its objectivity is open to serious question. Scheer made use of three broad categories as follows: drawings presumably inspired by observation of nature, those arising from personal experience, and those inspired by imagination or fancy. Examples of the first group include landscapes, houses, and the like. The second group includes pictures of baseball and other games, labelled or otherwise named pictures of known persons or animals, and so on. The third group includes decorative designs, fancy heads, etc. The implication that generalized patterns of interest may be inferred from evidence such as his is of doubtful validity.

Shaffer (**273**) compared the ability of public school children to interpret newspaper cartoons and found a steady improvement from Grade IV to the end of the high school period. This probably reflects an increasing acquaintance with public affairs quite as much as mental development.

✓ *The Psychology of Drawing: General Theoretical Material.* Among the more recent theoretical discussions of the developmental aspects of child art may be mentioned the work of Schaefer-Simmern (**263**), who bases his conclusions on the examination of the drawings of several hundred children. In its terminology at least, his position is unlike that of most other writers on the subject. He holds that a child's drawing is dependent upon a process which he calls "visual conceiving" which is based on perception rather than on abstract thinking, and which conforms to the principles of gestalt psychology. He regards visual conceiving as an emergent process, in which respect it resembles intelligence, but he maintains that it is a process in its own right, not a manifestation of intelligence, a form of language, or a medium for the expression of emotions. His comparisons with gestalt theory are, of course, not new. Similar statements have been made by many others, notably Hildreth (**127**) and by Bender (**26, 28**).

Schilder (**265**) objects to the use of the word "symbol" or "symbolic" as applied to the majority of children's drawings. He considers that such terms are appropriate when the sign or symbol is used as a means of conveying a meaning that is distinct from the manifest content. He maintains that most of the drawings made by children are overt attempts to depict reality; they become covert or symbolic only when the normal expression is blocked.

Cameron has advanced the remarkable contention that "in the normal person there is an arrest of development and often a decline during the age-period between eight and ten years" (**51**, p. 181). As evidence he cites the growth curve published by Peck (**241**), who found that scores on her drawing test improved most rapidly during the early years with negative acceleration becoming fairly marked after the age of seven or eight. This is quite in accordance with what has generally been found for most mental traits, although in this case it is possible that a more exacting method of scoring the test might have delayed the flattening of the curve to some extent.[3] Cameron's statement seems to neglect the leveling-off in growth of most aspects of ability and behavior incident to the approach of maturity. He ascribes this "decline" to the competing influences of group play and other social factors. There may be a lessening of interest in drawing at about that age though we have seen little evidence to support such a conclusion. Certainly the *ability* to draw does not decline but continues to improve until maturity, although changes are small after the periods of early and middle childhood have passed.

---

[3] The test used is described in a later section of this article.

Griffiths' study of imaginative expression in children (**112**) includes an interesting chapter on the early stages of drawing. She notes that after the scribbling stage has passed, the child is likely to center his attention upon the representation of a single object which he draws repeatedly until his work is reasonably satisfying to himself. Later he begins to add other objects pretty much at random; he is still engrossed with the problem of representing objects as such, not with making a synthesized picture. Only later does he begin to represent a theme as a whole. Fordham (**97**) has applied the psychoanalytic theories of Jung to the interpretation of child behavior. Two chapters of his book are devoted to drawing. He holds with Jung's theory of "archetypes" and believes that drawings, like dreams, show a mixture of outer and inner experience. Even the scribbles of the very young child have meaning, but this must be interpreted in terms of the content. Like Seeman (**271**) and Read (**254**), he attaches special significance to the circular movements known as the "whorl," which a number of people have observed to be characteristic of the early stages of drawing or scribbling. Lange (**153**) emphasizes the subjectivity of the art work of the preadolescent as compared with the relative objectivity and greater self-criticism characteristic of the postadolescent. Others have called attention to this point, and have noted that this change in attitude reduces the value of drawing as a means of personality diagnosis or at least renders such diagnosis much more difficult for older than for younger subjects.

The studies thus far described have depended in the main upon observation and description as methods of classifying their data. Quantitative results, when given, have consisted chiefly of simple percentages; in many cases only illustrative instances have been given to support the conclusions drawn. The investigators, for the most part, have been little concerned with questions of the relative merit of the drawings they examined; rather, they have sought to ascertain how children draw, in what ways their drawings change with advancing age, and, most of all, why they draw as they do. Others have been more interested in the use of drawing to throw light on problems of general psychology, particularly those having to do with visual perception and apprehension, and the changes in these processes that take place with the passage of time.

There are, however, many teachers and supervisors of art who recognize the fallibility of human judgment and seek to find better and more objective ways of measuring the quality of the art products of children. This search has led to the development of a number of devices designed not only for measuring achievement in art, but also for predicting artistic aptitude and for appraising art appreciation. Others whose interest centers primarily about the field of measurement have studied the reliability and validity of some of the scales already published or have attempted to ascertain whether or not their usefulness can be extended to fields other than those for which they were originally designed. All this has resulted in a fairly large body of literature devoted primarily to methodology, rather than to questions of development and behavior, for which the answers will obviously depend on the accuracy and freedom from bias of the methods used for their solution.

METHODS USED FOR THE STUDY OF CHILDREN'S DRAWINGS

Methodological studies of art products may be classed under several general headings as follows: (1) those designed chiefly for the appraisal of the artistic merit of a piece of work or the degree of artistic appreciation possessed by a given subject. These include chiefly objective tests, formalized rating scales, and objectively defined methods of classification, as well as some simple types of apparatus such as that used by Barnhart (**19**) which was described in the preceding section of this paper; (2) studies of the Draw-a-Man technique; and (3) new methods of observing, recording, classifying, and evaluating children's work, and of scaling art products in at least a reasonably objective manner with a view to learning something about the personality or abilities of the young artists. Only a few of the "projective" methods have been included in this group, since these studies for the most part have made use of intuitive rather than objectively defined procedures for their interpretations. The greater number of these articles are discussed later in this paper.

Scales and Tests. Among the tests and scales designed for measuring achievement in art may be mentioned the second of the two Kline and Carey scales (**143**) which is intended to measure the quality of design shown in posters, decorative borders, and similar compositions.[4] The method of deriving the scale values was the same as that used by Thorndike in 1913. Doms (**67**) also devised a scale of this kind, but this we have not seen. The McAdory Art Test (**178**), which requires the subject to judge the relative merit of the four plates comprised in each of a series of seventy-two items, was standardized on the basis of the ratings given by one hundred judges, all of whom were recognized artists, art teachers, or art critics. Although intended for use over the entire range from first grade through college, it is not well suited to the younger ages. Voss (306) has prepared a modified form of the test intended for use in grades below the fourth. The major difficulty with the McAdory test, however, as a number of its critics have pointed out, lies in the fact that so many of the pictured items have to do with costumes, automobiles, utensils, and the like for which styles change rapidly with the passage of time. Although the practical and homely nature of such a test appeals to many, it is questionable whether the basic principles of good design are likely to be recognized under so heavy an overlay of circumstantial and temporary factors. The Meier-Seashore Art Judgment Test (200) avoids this difficulty by utilizing pairs of sketches, one of which conforms to some well-recognized principle of composition or style, while the other violates it. By utilizing copies of the works of the great masters which have already stood the test of time, the validity of the test is, in a sense, guaranteed. The test can be used with children as young as the junior high school age, but is better suited to the senior high school or the college level. Several other tests suitable for the measurement of art appreciation or achievement at the higher age levels have also been devised, but since they involve little that is new in the way of method are not mentioned here.

---

[4] The first of the Kline and Carey scales which was designed for the measurement of free-hand drawing was described in an earlier article (**105**).

Lark-Horovitz and her associates at the Cleveland Museum of Art (157) have developed a series of samples illustrating various degrees of merit in different types of children's drawings. These samples do not constitute statistically graded scales; they are arranged only in presumable order of merit. They are designed primarily as an aid to teachers of art who are interested in grading the work of children. Peck's (241) scale, mentioned in a preceding section of this paper, was designed for the prediction of school success and is probably as good a measure of intelligence as of ability in drawing. The method is interesting. It requires the subject to copy photographs of common objects by means of a pencil. Scoring is in terms of points assigned for the representation of each of a number of specified parts. More conventional is the type of scaling used by Kerr (140) in her study of children's drawings of houses and by Tiebout (295) for the measurement of children's paintings, both of whom developed their scale values on the basis of differences equally often noticed by presumably competent judges. The Horn-Hellersberg Test (123) consists of twelve squares into which lines derived from test pictures by Carl Horn[5] are to be incorporated in such a way as to form a picture. The test may be used with adolescents. Hellersberg believes it to be a rather sensitive indicator of the subject's adjustment to reality, but the evidence for this contention is as yet incomplete.

The Draw-a-Man Technique. After a preliminary survey of methods (232), Oakley devised a measure of the drawings of a man made by adolescents (233). This scale may be looked upon as an extension of the Goodenough Draw-a-Man Test which appeared in 1926, designed for use only in the kindergarten and primary grades. An interesting feature of Oakley's scale, if it is proved to be valid, is a measure of egocentricity which is based upon the ratio of points most frequently scored by children of a given age to those less often scored. The idea is plausible but requires further confirmation. A number of M.A. theses done at Columbia University (58, 74, 79, 103, 164) have been devoted to the general topic of ascertaining the usefulness of the Goodenough test for subjects older than those for whom it was originally designed. These we have not seen. Berdie (**32**) reports a correlation of .62 between Goodenough and Wechsler IQ's with older adolescents of average and below average ability.

---

[5] The Horn Art Aptitude Inventory on which this test is based was devised primarily for use in art schools rather than for use with children in general.

Wintsch (**317**) reports an extensive study of children's drawings of the human figure and gives norms for French-Swiss children. He has also used the Fay test with various groups of European children, as have Fontes and others (**96**) and André Ray (**252**). The Fay test consists of asking a child to illustrate the situation, "A lady takes a walk and it rains." The subject has five minutes to complete the drawing. Scoring is by a system similar to Goodenough's. André Ray does not find the measure useful after age twelve; Wintsch suggests the age of nine. Ray (**253**) has also experimented with drawing piles of small wooden cubes from different angles as a test of mental development. The merit of this test, he believes, lies in supplying a concrete model, which reassures the timid child who may lack imagination. The author's confidence in children's ability to draw from a model is not supported by the findings of others that the presence or absence of a model makes very little difference to young children. Prudhomme's (**250**) more extensive and critical work is based on several thousand subjects from three-year-olds to the adult level, who drew human figures, animals, and geometric figures.

The reliability and validity of the Draw-a-Man Test has been the subject of a number of investigations, of which the most thoroughly controlled was carried out by McCarthy (**179**), who gave the test to 386 children in the third and fourth grades in New York City schools on two occasions with an interval of one week between the two administrations. Scoring was done by graduate students after a period of preliminary training. Each drawing was scored three times: twice by the same person and once by a different person. The correlation between self-scorings was +.94; between scorings by different persons, it was +.90. Self-correlation by the split-half method was +.89; by the retest method it was +.68.[6] Williams (**314**), who had one hundred drawings scored independently by five examiners, found intercorrelations between scores ranging from +.80 to +.96. The correlation with the Stanford-Binet was +.65. McElwee (**182**), working with subnormal children fourteen years of age, obtained a correlation of .72 with the 1916 Stanford-Binet. Yepson's (**324**) findings for feebleminded subjects indicate a retest correlation of approximately +.90, but a correlation with the Binet of only +.40.

McCurdy (**181**) obtained drawings of a man from fifty-six first-grade children on two occasions, separated by about three months. The two sets of scores derived by the Goodenough technique correlated +.69, and the mean IQ equivalents differed by a very small amount, statistically insignificant. Of particular interest is the collection of over one thousand drawings made by a boy subject in this study between the ages of two and one-half and seven years. This collection was made available by the parents when they learned of McCurdy's study. Fifty-six drawings of men were selected from this collection, the majority of the complete human figures available, excluding a number "because they were obviously intended as comic distortions or were obviously not done with the usual amount of care." The IQ scores for these drawings averaged 171 with an SD of 22. The two drawings obtained at school differ somewhat in IQ score and correspond roughly with values obtained from drawings made at home at about the same time. Likewise two Binet tests show a direction of IQ difference paralleling Goodenough IQ's obtained from drawings made at home at about the same periods. McCurdy concludes that variation in Goodenough scores rests upon fluctuation of ability which is more than "error of measurement," and that, in this instance, variability of the individual and variability of the group are of about the same magnitude. Unfortunately, individual and group variability are not comparable when mean performance differs so greatly. Only one of the fifty-eight IQ determinations for the boy fell below about the 97th percentile of general norms.

---

[6] The average of the three scorings was used as the basis.

A series of growth curves showing progress on the Draw-a-Man scale when the counted items or "points" were reduced to equally spaced units according to the Thurstone method of scaling was constructed by Williams (**316**), who found that the growth of children whose school progress had been rapid, average, or slow was very similar in pattern, though proceeding at different rates. All three curves show fairly marked negative acceleration.

A number of studies have dealt with the question of factors affecting success in drawing the human figure. Ames (**6**, **7**) found that children under the age of five depicted more details when adding to an incomplete figure than they did when told merely to draw a man, but that the reverse was true with children over five. However, the Gesell incomplete man, which was the model used, does not suggest the need of many details for its completion. A more sophisticated drawing might have yielded very different results. The *number* of parts added to the incomplete man (**6**) shows a more consistent and steady age relationship than do the data relating to specific parts added. This conclusion is in definite agreement with the basis of the Goodenough test.

Gridley (**111**), with four-year-olds as subjects, compared the results of nine prescribed drawing situations. The children were instructed: (1) merely to "draw a man," (2) "a little man," (3) "a big man," (4), (5), (6) repetitions of (1), (2) and (3) in that order, (7) a schematic drawing of a man from copy, (8) drawing a man from dictated instructions given one at a time as the children drew the different parts, and (9) drawing after the same instructions as were used in (8) had all been given. When scored by the Goodenough method the averages for all except Situation 8 were very similar, but, as might be expected, the dictation of what was to be drawn brought about an appreciable increase in the score, although neither the provision of a copy nor previous instruction had much effect.

Mott (**213**) used the Goodenough scale to compare the drawings made by fifty-eight children between the ages of four and seven years when the instructions given were simply to "to draw a man," with drawings made when the instructions were to draw a policeman, a farmer, or a cowboy. The majority of the children did better under the first type of instruction, apparently because in drawing the more distinctive figures their attention was diverted to the specific features of the costume instead of being centered on essential aspects of the human subject as such. In another study (**214**), Mott investigated the effect upon drawing scores of previous movements of specified parts of the body. Drawings were first made with only the general instruction to draw a man. Later the children were put through a series of exercises involving special parts of the body, verbalizing as they did so, e.g., "This is my head, I nod it." Drawings made immediately afterward showed that the exercised part was not only more likely to be shown but was drawn with more care for details. Two previous studies by the same author (**211, 212**) utilized drawings of the human figure done on three successive days with slightly different instructions on each day. The drawings were rated for five characteristics. Correlations with various other measures were of about the usual magnitude. The most interesting feature of these studies is that in each a fairly high correlation between ratings on the amount of activity shown in the drawings and those obtained on the Marston Scale for rating introversion-extroversion appears. Because of insufficient data concerning the character of the ratings and the age-range of the subjects, corroboration of the findings is needed.

Israelite (**137**) compared the relative difficulty of the items on the Goodenough scale for normal and mentally defective children of the same mental age. In general, the defectives surpassed the normals in respect to the number of details shown, while the normal children excelled on items involving the correct organization and proportion of the parts. Earl (**75**), who compared the drawings of adult defectives with those of normal children, found much the same thing to be true. Spoerl (**281, 282**), likewise, found that retarded children showed more details in their drawings than did normals who earned comparable total scores, but their sense of proportion was very poor. McElwee (**183**), who compared the profile drawings of normal and subnormal children, found that although the latter depicted more details than the former, the number of incongruities resulting from confusion between the full face and the profile position was much greater in the drawings made by the subnormal than in those made by the normal children. McHugh (**185, 186**) made an item analysis of the Goodenough scale with view to shortening it. He discovered that thirty of the fifty-one points in the scale correlated significantly with his criterion—the Stanford-Binet. His findings with respect to the abbreviated scale, however, are questionable, inasmuch as he used the same group of subjects to validate the new scale as were used for the original derivation. He also found that significant IQ gains were registered on a second drawing of a man after an interval of a few months and concluded that the drawing of a man is not free from the influence of experience.

Springer (**285**) made an item analysis of the Draw-a-Man scale to ascertain which items differentiate between maladjusted and well-adjusted boys as classified by the Haggerty-Olson-Wickman Scale. Of the 51 items included in the Goodenough scale, 15 showed critical ratios of 3.00 or higher. It is interesting to note that the pattern of the differences is the same as that found by those who have compared normal and retarded children; this in spite of the fact that the two groups were equated for total score on the drawing test. Berrien's (**33**) study, discussed later in this paper, attempts an item analysis to obtain a diagnostic measure of different types of behavior disorder in children. The possibility of using the Draw-a-Man Test as an indicator of maladjustment was also studied by Brill (**41**), who found that his maladjusted boys typically scored lower on this scale than they did on the Binet while the reverse was more often true of the well-adjusted. His attempt to devise a short form of the drawing test for the purpose of predicting maladjustment commits the too-familiar statistical error of using the same data for validating as for deriving the scale. It is time that prospective test-makers learned better than that!

Bender (**29**) used the Draw-a-Man Test with a small group of children suffering from post-encephalitic behavior disorders. In each of six reported cases, the IQ obtained from the drawing is far below that obtained on the

Binet. Bender's interpretation is that the cerebral damage had caused a considerable disruption of the body-image and suggests that so marked a discrepancy between the two scales may afford a useful hint as to the possibility of brain injury.

One of the best studies of the comparative performance of delinquent and non-delinquent subjects on the Goodenough test was made by Hinrichs (**130**). Unfortunately, his subjects (ages 9 to 18 years) were for the most part older than those for whom the scale was intended. Hinrichs attempted to remedy this difficulty by devising an upward extension of the scale but found, as have others, that the added section was neither as reliable nor as valid as the part designed for younger children. His subjects were 81 boys from an institution for delinquents, none with IQ's below 80. Several control groups from various types of environment were matched with the delinquents. In general, the study is well controlled and the statistical treatment is adequate. The chief findings are as follows: (1) somewhat lower mean IQ's on the Goodenough than on the Stanford for the delinquents but not for the control groups, even though chronological ages and home backgrounds were similar, (2) more "incongruities" in the drawings of the delinquents than in those made by the controls, (3) greater juvenility in choice of subjects and more "blood and thunder" characters among the drawings of the delinquents than in those of the controls, and (4) more stereotyped drawings with fewer indications of activity among those made by the delinquents.

When the evidence from these studies is put together, it appears that a Goodenough IQ markedly lower than that earned on the Binet may afford some indication of emotional or nervous instability or, it is just possible, of brain damage. However, inasmuch as the self-correlation of the Goodenough test is not high enough to warrant its use for exact comparisons, a finding of this kind cannot be regarded as more than one of the signs comprised in the total syndrome. If, in addition to the difference between the two test results, there are many incongruities or other unusual features in the drawing, the likelihood of personality disturbance is increased, especially if the Binet test shows the child to be of normal or near-normal intelligence.

New Methods for Studying Child Art. The series of studies carried out by Norman Meier and his students at the State University of Iowa suggests many useful ideas for the study of child art, even though few of them were carried far enough to warrant more than very tentative acceptance of the findings. Of the three monographs, only the first (**195**) deals specifically with the art of children; hence the others will not be considered here. The paper by Daniels (pp. 1–11) describes a method for studying the recognition of compositional balance by children of preschool age. Pairs of block designs, one of which was balanced in composition, the other unbalanced, were constructed within recesses in a wooden frame. The balanced model was sometimes on the right, sometimes on the left. The models were placed before the child, who was given a set of blocks and asked to build a duplicate of whichever one he wished. After he had finished he was told to look at the models again and choose the "nicer" one. Thirty-eight children were subjected to 185 experimental situations. Among the constructions, there were only 18 clear attempts to copy the unbalanced design. In only 31 choice situations were unbalanced compositions preferred. The author states on page 9 that since neither of these proportions is reliably greater than chance, the results indicate that "no child evinced a definite *preference* for the unbalanced design." This assertion obviously indicates lack of understanding of the null hypothesis. Jasper (pp. 12–25) made use of four devices for studying rhythm in graphic form: (a) a completion test in which the child was required to select from a number of alternatives the appropriate card to fill a space in a border made up of pictures of concrete objects, (b) a similar test in which abstract forms were substituted for the pictorial elements, (c) a dual choice test in which the two cards were used, one of which contained a rhythmic, the other an arhythmic arrangement of the same seven elements. (d) a construction test similar to that used by Daniels except that rhythm or its lack was the differentiating factor. In all four tests, regular progress with age was observed but other findings were of little significance.

Whorley (pp. 26–45), after a careful analysis of the literature on the subject, decided that the elements of compositional unity upon which authorities seemed best agreed are three in number, viz.: emphasis, balance, and fitness. A test was designed for each. Emphasis was studied by requiring the child subject to arrange four toy trees about a bird-bath. Balance or symmetry was studied by the arrangement of the trees about a central archway. Fitness was evaluated by the arrangement of furniture within a dollhouse. Scale values were obtained by utilizing the median value given by art experts to each of a series of photographed arrangements chosen to represent a roughly graded series. Scores were assigned by comparing the children's arrangements with the photographs as standards. The subjects were one hundred children from three to eight years of age and fifty adults. Little success was attained by the children under four; thereafter improvement occurred, but even among the adults, success, according to the scoring method employed, was far from universal. Whorley reports very high reliability coefficients for the method but inspection of the scatter diagrams shows that these figures are in error by a considerable amount.

Williams (pp. 46–50) has presented a brief report on a method of testing the sensitivity of children to color harmony. A doll-dressing procedure was used. Two identical sets of six small dolls were prepared. Each doll in a set was dressed in different colors. The dolls were handed to the subject one at a time, together with four scarfs of different colors. The child was asked to choose the scarf that he thought "would look nice with the doll's pretty dress." One of the scarfs was of a harmonizing color; the other three were chosen to give as poor an effect as possible. The experiment was repeated, using the second set of dolls. Williams' findings were combined with those of Walton (pp. 51–62), who used the same method with approximately six hundred subjects ranging in age from four years to the adult level. The pattern of change in score with age is curious. Both the mean and the median show a steady *decrease* up to the age of seven, after which improvement continues until the adult level is reached. It would appear that the ideas of color harmony prevalent among young children differ materially from those of the older persons in our culture. It would be interesting to try the same experiment with primitive groups who have had little contact with white civilization. Sex differences were small but consistently in favor of the girls. Only a very slight difference was found between the performance of a small group of children considered by their teachers to be gifted in art and that of another group judged to lack artistic ability. No relationship to measured intelligence was found.

Grippen (pp. 63–81) studied creative imagination in nine children by what she termed the "constant contact" method, which means only that the children were observed throughout the drawing period and were encouraged, by means of questions and comments, to talk about their work. Four of the subjects were considered talented in art; the other five had shown little interest in it. Apart from the expected differences in the quality of their drawings, the talented children showed more interest in their work; they talked chiefly about their drawings while the untalented were more likely to chatter on irrelevant subjects. Dow (pp. 82–94) found that artistically gifted children were less active on the playground than children with little talent, and that they spent more time playing with materials and less with other children. Rodgers (pp. 95–107) took approximately twenty-five photographs of the interior and exterior features of the homes of each member of two small groups of children considered to be talented in art and for two corresponding groups regarded as untalented. These photographs were projected on a screen and rated by a number of art experts as to the esthetic qualities of the homes depicted. Although for both of the two sets of data compared the advantage lay with the gifted group, the cases were too few and exceptions too many to justify conclusions. The final study in this group is by Tiebout (pp. 108–133). She found that artistically gifted children were superior to the non-gifted on tests of observation and visual memory but not on motor tests. This study, incidentally, is the only one in the series in which the statistical methods appropriate for small samples have been used. The monograph is interesting because of the variety of the problems and the originality of the methods used for their solution, but for the most part the treatment of the data is poor.

The superiority of objective over subjective methods for rating the quality of drawings was demonstrated by Brooks (42), who had thirty drawings of a house and thirty of a boy rated independently by each of twenty-one college students, first on the basis of unaided judgment, then by the use of the Thorndike Scale, and finally by the Kline-Carey Scale. Variability of judgment was only about half as great when a scale was used. The Thorndike gave slightly better results than the Kline-Carey. Similar results were obtained by Lark-Horovitz (156) who lays particular stress on the large number of judges needed for even moderate reliability when the subjective method is used.

"Free" or "spontaneous" drawings are often not so spontaneous as the name implies. In an attempt to find drawings which had certainly been done without direct suggestion from adults, Bender (26) studied a large number of the sidewalk drawings made by children. She was impressed by the many geometrical designs and by the frequent omission of details. The whole rather than the parts seemed to be the dominating principle in making these drawings. Although, as she points out, this observation is in accordance with gestalt psychology, the rough surface of the walk and the coarseness of the medium employed (usually chalk), must also be considered. In a somewhat related study Hurlock (134) collected 1,451 drawings from such unconventional sources as waste baskets, class notebooks, and the like in a number of high schools and colleges. Classification by subjects gave the following results: Decorated words or letters constituted 27.3 per cent of the total; caricatures, 22.8 per cent; human figures, 17.2 per cent. The remainder were divided among more than twenty other classes, with no single topic accounting for more than 8 per cent. Of the human figures well over half were drawings of men; the remainder represented girls or women. Not one of the drawings represented a baby or a child. In view of the age of the subjects, this is an interesting point.

Both Dallinger (61) and Spielrein (280) have pointed out the relationship between the child's growing awareness of his own body parts and his depiction of these parts in his drawings. Hartenstein (119) and Fontes (95) compared the roles of the literal and the fanciful in children's drawings; the former using German, the latter Portuguese children as subjects. Fontes considered only the proportion of "supernatural" objects such as angels, fairies, the Infant Jesus, etc. The numbers were small at all ages but were many times more frequent among drawings made by girls than among those made by boys. Hartenstein, using a more liberal criterion, found that boys were about twice as likely as girls to depict objects actually seen. She noted also the effect of siblings and associates of similar age upon the character of children's drawings. This she believed to be much greater than that of parents or teachers. She observed that the drawings of girls with brothers but no sisters were likely to be modified in the direction of masculinity and that the converse was true with boys whose only siblings were girls.

Anastasi and Foley (10) found many resemblances to the art of children in the work of non-professional artists who paint merely for their own pleasure. White and Johnson (312) report that children of nursery school age were more successful in interpreting drawings made by their contemporaries than in naming the subjects shown in modern paintings, particularly those of the impressionist school. Many people will not find this surprising!

An experiment by Kröber (150) would bear repetition as a developmental test, using subjects with a wider age-range. Kröber had over four thousand subjects between the ages of thirteen and sixteen years attempt to draw from memory a picture of a formboard into which they had previously fitted eighteen simple geometrical figures—a circle, a square, a star, etc. Half of the subjects made the drawing immediately after fitting the forms in place; the remainder, after a lapse of two hours. No age differences were noted within the range covered, but boys reproduced more figures than girls though the sex difference was not large. The more complex the figure, the more likely it was to be modified in accordance with its expressive value for the subject.

Family resemblances in drawing were studied by Kretschmann (148), who had ten judges attempt to match the drawings of parents and children by a method similar to that of "blind analysis." Success was considerably better than chance. Various cues having to do with similarity of style were reported. Seeman and Saudek (272) studied the handwriting and drawing of three pairs of identical twins. Their major interest was in the handwriting, which is discussed at some length. Specimen drawings are reproduced with little comment but the similarities in style are apparent. It is hardly necessary to point out that both these findings can be accounted for quite as readily by mutual imitation and environmental motivation as upon the assumption of the inheritance of some special talent or behavioral tendency. Meloun (201), in common with several others, found a low positive correlation between skill in drawing and quality of handwriting. This is quite to be expected.

Factor analysis has been applied to scores based on drawing tests or products by Findley (90) and by Koga and Naka (146). The factors identified in both cases appear to involve perception and memory. Neither study discusses a motor performance factor.

THE INFLUENCE OF WORLD WAR II ON CHILDREN'S DRAWINGS

It has been assumed by many that emotional stress incident to war—the distress due to the breaking up of families and more especially the terror experienced by those directly exposed to conflict—would be mirrored in children's drawings. Perhaps this is the case, but if so, those who have attempted to study the matter have not been sufficiently skilled in the language of child art to read what is there. In most quantitative studies the analysis has been made rather superficially in terms of the percentage of drawings which overtly depict war themes. To the present writers it would seem more reasonable to expect that children genuinely and deeply disturbed by such experiences would be likely to express their feelings covertly rather than overtly, and that only painstaking search would reveal the nature of these signs to an outside observer.

Direct effect of the war upon children's drawings does not appear to be large. Hildreth (**128**) reports an increase from 14 to 22 per cent of war themes in the drawings of kindergarten and primary school children from 1942 to 1943. This may or may not indicate a genuine change, inasmuch as the numbers of children included were small, and boys, who made up a larger proportion of the 1943 than of the 1942 sample, contributed most of the war themes. Naumburg (**224**) presents a descriptive account of the war themes drawn by several children who had been brought to a behavior clinic. She expresses the opinion that the continued drawing of such pictures has a therapeutic value in providing an outlet for suppressed feelings of aggression and in clarifying emotional conflicts. Bender and Wolfson (**31**) describe the nautical drawings of children before and after the outbreak of hostilities. The latter series depicted more detail, especially insignia, and rather more war and combat themes, but the differences are small.

The most complete study of this topic that has come to our attention is by Baumgarten and Tramer (**21**). The subjects were 272 Serbian refugee children who had been evacuated to a camp in Switzerland. Their age-range was from five to twelve years. All had experienced war at first hand. Many had seen their homes destroyed by bombs; a large proportion had lost one or both parents. Nevertheless, when free choice of subjects was allowed, few children elected war themes. Those drawn were for the most part unemotional in character—tanks, soldiers marching, airplanes. More than 90 per cent of the free drawings represented home scenes, flowers, birds, or people engaged in civilian occupations. Quite possibly these drawings are analogous to the "wish-fulfillment" aspect of dreams pointed out by Freud; they represent the children's desire to forget the painful experiences through which they had recently passed and to experience peace and security once more.

DRAWINGS BY THE DEAF AND BY THE BLIND

Upon the assumption that the elimination of one sensory avenue may give rise to a compensatory superiority in others, a number of persons have examined the drawings of the deaf. Three of these studies have utilized the Draw-a-Man Test, assuming that this test not only may provide more objective evidence of compensatory mechanisms than is given by subjective judgment, but also may constitute a better index of the intellectual level of the deaf than is given by tests of a more verbal character. The three studies agree in finding mean IQ's for the deaf that fall somewhat below the mean for the hearing. The averages are as follows: Peterson and Williams (**243**), 80; Shirley and Goodenough (**275**), 88; Springer (**284**), 96. The first two studies utilized institutional cases. The last included day school children enrolled in special classes for the deaf in New York City, who probably represent a higher socio-economic level, on the average, than children who have been placed in institutions. Thiel (**291**) on the basis of an examination of approximately two thousand drawings made by children in schools for the deaf came to the conclusion that the development of drawing among the deaf parallels that found for the hearing but that progress is slower. Pintner (**246**), using the McAdory Art Appreciation Test, found that deaf subjects between the ages of eleven and twenty-one years performed as well but no better than hearing subjects of corresponding age. Although the deaf girls did slightly better than the boys they showed no superiority over hearing subjects. All this is in direct opposition to earlier studies in which more subjective and possibly more biased methods of evaluation were used.[7]

By far the most important study of the art work of the blind and weak-sighted is found in Lowenfeld's *The Nature of Creative Activity* (**172**), which will be discussed in the section on projective methods. Cook (**59**) has described methods for training the blind in artistic expression. He expresses the belief that many blind are capable of art work of exceptional merit since they are not hampered by a knowledge of the non-essential details that confuse the seeing person. The possibilities of finger painting in treating the emotional problems of the blind have been noted by Napoli and Harris (**221**); a number of case studies show how this procedure can serve as a means of mental catharsis whereby the blind patient may objectify the feelings of resentment and aggression that result from the frustrations due to his condition.

---

[7] See Goodenough (**105**, p. 276).

THE DRAWINGS OF CHILDREN IN DIFFERENT CULTURE-GROUPS

With the relegation of the theory of recapitulation to the limbo of outworn beliefs, comparisons of the drawings of modern children with those of prehistoric man have well-nigh disappeared from the literature. Prehistoric art is now studied for its own sake and for the light that it can shed upon the culture of the times. A few of the earlier studies in the period covered by this review have discussed the similarities between child art and prehistoric art more or less briefly. Eng's critical discussion (**83**) is enlivened by many excellent reproductions which illustrate the points of similarity very well. Luquet (**175, 177**) has elaborated and extended his earlier work on this topic and Boas (**39**) has included a few notes on the art of modern children in his extensive anthropological report on primitive art.

A number of people, however, have been interested in the art products of the children of primitive groups. From drawings originally collected in 1911 among certain very primitive Indian tribes living in the high plateaus of Bolivia, Rouma (**259**) made a number of interesting observations. He states that the drawings made by adults in these groups were hardly distinguishable from those made by European children under the age of six. The general scheme of the drawings was much the same and they exhibited the same kind of confusions. Many of the Indians had difficulty in orienting the drawing on the paper. "Up" to them meant toward the sky, not toward the top of the sheet. A few of the Indians became interested in drawing and improved with practice, but none advanced beyond the level shown by European children at the age of 12. Du Bois (**71**), in her study of Alorese, reports 99 free drawing performances, including over one thousand figures, of 33 boys and 22 girls, between 6 and 16 years of age. The human figure appears in only 7 per cent of the boys' drawings and in only 1 per cent of the girls'. Plants, animals, and buildings account for 64 per cent of the boys' figures, and plants, tools (including tattoo designs and combs), and buildings make up 84 per cent of the girls' figures. Du Bois remarks that children's drawings, ordinarily made as a play activity, have no relation to Alorese adult art, which is primarily stereotyped and decorative. When specifically asked to portray the human figure, one group of children produced results which were well below the Goodenough age norm, but the drawings of a second group ranked at the standard. The author prefers to speak of differences in "aptitude, adaptation, and the development of skills" (**71**, p. 578) rather than of comparative intelligence. As in other primitive groups, the children showed no consistent directionality of drawing and no attempt to portray space. Appia (**17**) obtained drawings of the human figure from Negro children in French East Africa. These drawings showed the developmental accretions commonly found in European groups and were quite comparable in general with drawings made by European children of similar age. If anything, the Negro children showed more creative imagination than Europeans. The children drew typically from right to left, a finding not in character with the directionality noted in certain studies discussed elsewhere in this paper. Steggerda (**287**) obtained drawings of a man from 48 full-blood Maya and 330 Maya-Spanish children. Their performance was appreciably below the work of Navaho children and the Goodenough norms, age for age.

Paget (**237**) who studied more than sixty thousand drawings made by children from various parts of Africa, India, China, the South Sea Islands, and elsewhere, does not agree with the view that the difference between the art of primitive children and that of children reared under the influence of Western culture is in the main a matter of *level* rather than *manner* of development. He accepts the two general principles with which most investigators agree, viz.: that children draw not to depict nature but to tell what interests them; and that because they attempt scenes and objects far beyond their technical skill, they are forced to invent their own symbols or devices for indicating them. The second point has been particularly stressed by Fallaize (**88**) and by Naber (**217**). Not all groups adopt the same symbols, but there is a fairly high degree of similarity in those used by any particular group. Paget states (p. 138) that "however they may have arisen, distinctly local modes of representation exist among untaught children and these modes are often peculiar and unconnected with adult tradition." Among these specialized features he noted that 60 to 80 per cent of Maori children begin to draw the human figure in profile between the ages of six to eight years, whereas the average American or European child does not make the shift until around the age of nine. "Mixed profiles" also are rare among the Maoris but are very common among the West Africans. The "hourglass trunk" is also common among several widely separated groups, but is practically never found among others. The nose on the full-face human figure shows the widest variety of symbolizations from group to group. Paget is not inclined to attribute such localized modes of representation to adult teaching or example. He looks upon them rather as traditions handed down from one generation of children to another in much the same way as children's games.

Schubert (**270**) was so impressed by the naturalistic qualities of the drawings of Orotchen children who up to the three months before her study, had never seen a picture or handled a pencil that she was inclined to believe they possessed unusual qualities of eidetic imagery and that the remarkable excellence of their drawings resulted from tracing around the eidetic image as it was projected on the paper. This explanation is questionable, however, since Loch (**170**) and a number of other investigators have shown that the eidetic image can rarely, if ever, be reproduced graphically, even though it can be described with remarkable fidelity.

Of the descriptive studies comparing child art in modern civilized countries, by far the most elaborate is a report by Anastasi and Foley (**8**) on some six hundred drawings from children under the age of twelve. These drawings were displayed at Radio City, New York, in an international exhibition of child art. Forty-two different countries were represented. The study is severely criticized by Spoerl (**283**), who points out that (a) the small number of cases from many of the countries made it difficult to take chronological age into account when comparing the drawings, even though age differences between the child artists were frequently great, (b) mental ages of the children were not known, and (c) many of the differences ascribed by these authors to nationality correspond very closely to the age progression shown in the work of children in this country. An even more serious question not noted by Spoerl arises from the manner in which the drawings were selected. They were chosen by art teachers or art critics in the country of origin, and thus necessarily reflect whatever differences may exist from one country to another in respect to standards of what is good art. By no stretch of the imagination can they be looked upon as representative samples even of the best in child art from the countries in question.

A second study by Anastasi and Foley (**9**) describes the animal drawings of Indian children attending a government school in British Columbia. The authors state that both the subject matter and the technique of the drawings reflect special cultural and experiential factors rather than age differences or developmental trends, although they present few data. Inasmuch as more than half of the subjects were well over twelve years of age, developmental differences would, in any case, be slight, but the age factor should not be overlooked in interpreting this statement. Several illustrations are presented to show the highly stylized patterns used, but the age of the subjects is not stated.

The Draw-a-Man Test has been used in a number of comparative studies of national or racial groups. Manuel and Hughes (**190**) gave this test to two presumably comparable groups of Mexican and non-Mexican children in Texas and found a mean difference in IQ of approximately 10 points in favor of the latter. Eells (78) gave both the Draw-a-Man Test and a modified form of the Stanford-Binet to several native groups in Alaska. All groups earned mean IQ's in the low 90's on the drawing; no group averaged above 80 on the Stanford. Although differences are small, the rank-order of the various groups is identical on the two tests. Menzel (**202**) gave the Goodenough test to approximately two thousand school children in East India. The scores earned were far below those found for other national-racial groups. This was especially true of those earned by girls over the age of ten years—a fact which may be the result of the seclusion of girls as they approach puberty. Datta's (**63**) work with Bengali children revealed age progression, but did not indicate direct comparability with European results. The Draw-a-Man Test has been translated into Chinese by Chang (**55**) and standardized for Chinese children by Hsiao (**132**), with over four thousand elementary school children as subjects. Huang (**133**) is impressed by the similarity of developmental sequences in the drawings of young Chinese children to those found in the work of children of Western cultures, but it would be surprising if there were no differences between the American and Chinese standards for this scale.

In addition to the study by Eells mentioned in a preceding paragraph, the Goodenough scale has been used with at least four groups of American Indians. The first of these, by Telford (**290**), was carried out on a government reservation in North Dakota; the subjects were chiefly Chippewas and Sioux. Both full blood and mixed blood cases were included and the group as a whole earned an average IQ of 88. Telford was chiefly interested in repeating the studies carried out with American Negroes, which have, in most cases, shown a positive correlation between the level of intelligence and the degree of white blood. This was not the case for the Indians, however; as a matter of fact, a small but unreliable tendency in the opposite direction was noted. MacGregor (**184**) reports a study of Sioux in two South Dakota communities. Children in the one with more "white influence" averaged 102 IQ points and the other averaged 113.6. To the geneticist, such findings might indicate that interbreeding between Indians and whites has been carried on chiefly by the less intelligent members of one or both races.

Dennis (**64**) studied Hopi children, a group among whom graphic art has long held an important place. However, it is practiced chiefly by males. He found a marked sex difference in performance on the Draw-a-Man Test. The mean IQ earned by the boys was 117; that earned by the girls was 100. Breakdown of the figures according to age showed a fairly regular tendency for the boys to increase and the girls to decrease in relative standing as age progressed. A similar but more extensive study by Havighurst and his associates (**122**), which involved children from six different Indian tribes, showed a similar sex difference in all tribes where art work is chiefly the responsibility of the males. Russell (**261**), however, found that among Zuni children, culture patterns were less noticeable in the results of the Draw-a-Man Test than in drawings made when free choice of subject was permitted. This, of course, is by no means in opposition to the findings of Dennis and Havighurst. The tests reported by Havighurst are discussed somewhat more extensively in relation to anthropological findings in several monographs (**138**, **160**, **184**, **292**). The McAdory Art Test was given to Mexican children by Stolz and Manuel (**288**) and to Navaho Indian children by Steggerda (**286**). The Mexican children were also given the Meier Art Test. Both investigators failed to establish any positive relationship whatever between test scores and proficiency in the various arts and crafts practiced by these groups. Vydra (**307**) worked out a series of ratings for judging the artistic merit of drawing. He applied the method to compare the quality of drawings made by the children of different racial stocks in several elementary schools in Slovakia. The rank order was found to be as follows: Slovaks, Hungarians, Germans, Jews, and Gypsies. In America, Merry (**203**) could find no evidence of an unusually large number of artistically gifted children among those of Jewish, Negro, Italian, or "native white" stock. A brief account of the drawings of young children in Soviet Russia has been given by Labunskaya (**151**), who takes the position that child art reflects the social character of the time and circumstances in which it is produced and, as such, is worthy of careful study. He describes a number of different types of research said to be in progress but gives no results.

Developmental changes in the work of Brazilian and Bolivian children when free choice is permitted were studied by Rabello (**251**). As is usual in these studies, variability at all ages was large and there was much overlapping from age to age. The findings do not appear to differ greatly from those of children in other countries. In their study of the free-choice drawings of Serbian refugee children, Baumgarten and Tramer (**21**), however, noted an interesting relationship between the sex of the child artist and the apparent sex of the human figures drawn. At all ages, boys drew chiefly males. Girls under the age of ten drew male figures predominately. Inasmuch as the camp did not admit children over twelve years, the study yields no data on a shift toward a preference for drawing the opposite sex in boys as they approach puberty. As far as we are aware, the point has not been considered by other investigators.

Using Abel's (**1**) method described elsewhere in this article, Taylor (**289**) studied the drawings of young college students in Central India. He found a much smaller percentage of drawings belonging to what Abel termed the "schizophrenic type" than she reported. Of the various castes and religious groups included, Taylor noted that the Hindus and Jains showed less "unity" in their drawings than the other groups. Cultural influences seemed to persist even after several years of common school training which is similar for all sects. However, his numbers are in many cases too small to make the comparisons between subgroups very reliable.

Considering all the evidence thus far described, it is clear that cultural differences do appear to a greater or less extent in the drawings of children. The present writers would like to express the opinion that the search for a culture-free test, whether of intelligence, artistic ability, personal-social characteristics, or any other measurable trait is illusory, and that the naive assumption that the mere freedom from verbal requirements renders a test equally suitable for all groups is no longer tenable.[8] The studies by Dennis and by Havighurst point the moral very clearly as far as the Draw-a-Man Test is concerned.

There has long been an emphasis upon the need for higher self-correlations and larger correlations with an accepted criterion in a test designed for individual diagnosis than in one intended only for comparisons between groups. We also believe this emphasis has led many psychologists to overlook an equally important principle, viz.; that for valid group comparisons a test must be equally *suitable* for all the groups concerned. The determination of "suitability" is not an easy matter but it cannot safely be ignored. It requires the researcher to set up hypotheses and to test them by carefully conducted experiments. An item analysis will often suggest in what direction sources of cultural bias may lie, for performance on a single item is an easier fact to interpret from this standpoint than is total score on a many-itemed test designed to measure some general and, often, not-too-well defined trait.

We may illustrate our point by a further consideration of the findings for the Goodenough Draw-a-Man Test. Repeated studies have shown that when used with children of reasonably similar cultural backgrounds who are equally motivated to do well, the test is serviceable as a crude measure of "general intelligence," although the moderate self-correlations and correlations with outside criteria make it clear that it cannot serve as a satisfactory substitute for individual tests of the Binet type. From such facts, one might conclude that the test is better suited for ranking individuals than for ranking groups in order of estimated intelligence, provided that the individuals studied are comparatively homogeneous in respect to other factors which affect the results. Thus, the rule that a test of only moderate reliability and validity is better suited to the study of group differences than to the identification of differences among the individuals comprising any group can be accepted only with the reservation that the test is as well suited to the interests and experiences of one group as to those of the others.

---

[8] The early study by Goodenough (**104**) which reports the differences in standing on the Draw-a-Man Test by the children of immigrants to the United States from various European countries is certainly no exception to this rule. The writer hereby apologizes for it!

Special Artistic Talent

For valid diagnosis of talent in art, some reasonably objective criterion is obviously needed. A number of the studies mentioned in the preceding sections of this paper have a bearing on this question, particularly those dealing with tests and measurements. In this section we have included only those studies in which the aim was primarily to identify the artistically gifted, rather than to measure the extent of art talent in the generality. A few studies describing the characteristics or background of persons talented in art have also been included. Among the latter, the studies by Grippen (**113**), Dow (**68, 69**), Rodgers (**257**), and Tiebout (**293, 294**), described in an earlier section, should be noted.

Criteria for the identification of the artistically gifted are discussed by Cane (**52**), who describes four characteristics she has found useful: (1) the child's use of his own body (rhythm of movement, posture, etc.); (2) the fertility of his imagination and the quality of his feeling; (3) his ability to organize as shown in his portrayal of form and design, and (4) his "spirit or emanation of life" which necessarily results from the interrelations of the first three. Schmidl-Waehner (**268**) insists that formal criteria such as the size of the drawing, its position on the paper, the types of lines used, and the like are more revealing than content. McCloy's study (**180**) was concerned with the possible effect of lack of manual skill upon manifestations of talent in art. He devised an interesting apparatus in which the subject could achieve various compositional effects by manipulating colored lights directed upon a group of clay figures before a suitable background. As in many such studies, the lack of a suitable criterion for evaluating the results proved to be a major difficulty. Knauber (**144**) studied the art abilities of young children from the nursery school to the third grade and came to the decision that not more than one child in thirteen shows definite talent in art. Subjective judgment, however, was the only criterion used in rating the drawings.

Munro, Lark-Horovitz, and Barnhart (**215**) do not agree with the oft-expressed opinion that the truly gifted in art differ from the generality in possessing special characteristics or combinations of characteristics. According to their studies at the Cleveland Museum of Art, the gifted differ from those with little talent in degree, but not in kind. Age for age, talented children are more advanced in artistic expression. Their development of skill is more rapid and continues long after that of the ordinary child has ceased. They exhibit a stronger interest in visual experience and a correspondingly greater tendency to express themselves in visual terms, even attempting to express non-visual experience by means of graphic symbols. These writers describe a series of experiments and present age norms in terms of which differing levels of talent may, they think, be distinguished.

A number of persons have examined the thesis that artistically gifted persons are more likely than the generality to be morbid, emotionally unstable, eccentric in their habits, or even psychotic. Uniformly, the findings have failed to confirm this belief. Two examples may be cited. Carroll (**53**) evaluated the art ability of older adolescents, using the McAdory and Meier-Seashore tests and quality ratings by experts of art products. He found no relationship whatever to measured introversion, submissiveness, or emotional instability, or to judged general personality disorders. Farnsworth (**89**) had elementary school teachers rate their pupils on a number of carefully defined personality traits as well as for their ability in several types of fine arts including drawing and painting. No relationship whatever was found between rated ability in art and undesirable personality traits. If anything, the reverse tendency was indicated.

Low positive correlations between teachers' ratings on the personality traits of their pupils and ratings on artistic talent have been reported by Flemming (**93**). Cattell (**54**) has developed ratings on 35 qualities, selected to represent the totality of personality. Eight of these correlated significantly (though actual values were low) with artistic drawing ability. Subjects were about 180 young men. Generalizing from the specific findings, Cattell states (p. 140):

> The agreement of this syndrome with the popular stereotype of the earthy, gay, spirited, unstable, careless, Bohemian artist is very striking, yet it must be emphasized that at the time the ratings were made, no one had any idea that the purpose was to correlate them with artistic ability.

His criterion was a judgment by skilled art teachers of the drawing ability shown in work samples by his subjects. The ratings of personality were obtained from an entirely different group of judges. Cattell speculates that a person inherits the appropriate temperamental tendencies; he acquires drawing ability as such on the basis of interests which his temperamental qualities generate.

Case studies of the development and behavior of children who were judged to have marked talent in art or of the early work of those who, as adults, achieved fame as artists, are fairly numerous but few of the reports have much psychological interest. One of the best is by Beck (**23**), who describes the early development of Romano Dazzi and a number of other child artists, together with an analysis of the factors that seemed most influential in turning their attention toward art as a medium of self-expression and self-development. Meier (**197**) has reported the case of a boy, blind from birth, who at the age of seven, was given partial vision by means of an operation. Almost from the beginning, this boy showed exceptional artistic talent, particularly with respect to design and composition and the use of color. In view of the previous complete absence of visual experience or of special training in art, Meier looks upon this case as affording rather crucial evidence for the essentially innate character of artistic aptitude. In the third volume of his *Studies in the Psychology of Art* (198), Meier lists six factors, all of which he considers essential aspects of artistic ability. These are as follows: manual skill, energy output and perseveration, aesthetic intelligence, perceptual facility, creative imagination, and aesthetic judgment. He adds (p. 158):

> It is the thesis of the writer that the person with these six factors *can* bring this end [extraordinary artistic accomplishment] about and that the person without these *cannot* bring this about to any great degree. The position is, therefore, a deterministic point of view—deterministic in the sense that certain neurophysical and developmental factors seem to be normally a *pre-condition* for the rest of the total development, and that these predisposing conditions are not present equally in all persons, nor, if absent, can they be established. . . . The unique contribution which these fifteen years of investigation have made is the knowledge that all these factors are in a peculiar and unique manner *interlinked*, and the entire dynamic process is a closely knit, interdependent, and evolving process.

Very little seems to have been done in the way of developing objective methods for the study of the interlinkage of factors upon which Meier lays so much stress. In spite of the many who have devoted years of patient research to the problem, the question of the distinction between real artistry and good craftsmanship, if a difference exists, is still unsolved. The main difficulty, as we have said before, appears to be the selection of a criterion. We still ask, with Kipling, "It's pretty, but is it art?"

THE RELATION OF SPECIAL APTITUDE IN ART TO GENERAL INTELLIGENCE

Even before the publication of the Goodenough Draw-a-Man Test in 1926, most of those who had investigated mental development had come to realize that indications of a child's mental level may be observed in his drawings. This fact seems now too well established to require further consideration here. In this section, we shall mention only briefly a few recent studies dealing with the relation between artistic aptitude as such and general intelligence. For the most part, these studies have confirmed earlier findings that both subjective judgments of the artistic merit of drawings and scores on various art aptitude or achievement tests and scales show positive but low correlations with scores on standard tests of intelligence. Studies by Tiebout and Meier (296), Lewerenz (165), and others are examples. The unusually high correlations reported by Bird (37) and by Miller (205) are due to the wide range of ages included in their computations. Had age been controlled, the figures would unquestionably have been much lower.

The fact that a low positive correlation gives leeway for many exceptions to the general rule, some of which may be quite marked, is illustrated by a case reported by Vaughn and Hoose (**301**). A boy of fourteen whose Binet IQ by repeated tests had fallen consistently in the 60's, was found to have exceptional ability in copying drawings both directly and from memory. He also ranked very high on tests of art judgment and art appreciation and on tests of visual imagery. His original drawings also ranked above average in artistic quality. His school records were very poor, both in formal subject matter and in conduct. A change to a program making greater use of his special talent brought about marked improvement along both lines.

Abel and Sill (**2**) made an interesting comparison of normal and feebleminded subjects using a plan similar to some of the experiments of Lewin. Papers on which a 4-in. square had been drawn were given to over four hundred adolescents, who were told to divide it into smaller squares. Children with IQ's above 80 most often utilized cross-sectioning; those with IQ's below 80 were more likely to draw each square separately in a series of "nests."

Hurlock and Thomson (**135**) studied the relation to age and intelligence of seven aspects of drawing. They present the following conclusions: (1) The tendency to perceive the specific rather than the general increases with age; (2) The tendency to perceive background and color placement increases with age and, to some extent, with intelligence; (3) The tendency to perceive associated objects and design increases with age but otherwise shows no relation to intelligence; (4) The tendency to perceive details increases with age and also with intelligence; (5) Accuracy of perception is positively related both to age and to intelligence; (6) Artistic expressiveness as such shows little relationship either to age or to intelligence. These findings are in general agreement with the statement made in the first paragraph of this section, that the correlation between artistic ability and general intelligence is positive but low. In this connection, it may be noted that the estimated childhood IQ's of the famous artists studied by Cox,[9] although much above the generality, were still below the average of most of the other classes of historical geniuses studied. Only the soldiers ranked lower.

ART EDUCATION FOR CHILDREN

Inasmuch as most of the studies in this area are of interest to teachers and art supervisors rather than to psychologists, only brief mention of a few of them will be made here. By far the most complete survey of art education is presented in the *Fortieth Yearbook of the National Society for the Study of Education* (**222**). Other reviews of methods used in

---

[9] Cox, Catharine M. The early mental traits of three hundred geniuses. In L. M. Terman (Ed.), *Genetic Studies of Genius*, Vol. II. Stanford University, Calif.: Stanford Univ. Press, 1926.

Bureau Ednl. Psy. Research
DAVID HARE TRAINING COLLEGE
Dated

the teaching of art have been prepared by Garrison (**99**), Tomlinson (**298**), and Liese (**168**). Some discussion of instructional methods in art also appear in many curriculum studies.

Writers on the developmental aspects of children's drawings have generally expressed the belief that free expression is better than rigid conformity to rules and principles as a method of cultivating art talent. Even greater emphasis on this point is given by those who stress the therapeutic value of child art. This principle has been particularly stressed (though with differing expected outcomes) by Bird (**37**), Dumont (**72**), Foley and Anastasi (**94**), Lowenfeld (**174**), Read (**254**), Schaefer-Simmern (**263**), and Viola (**303**). Kerr (**141**), however, adds that freedom of expression must be cultivated, not left to spontaneity. A somewhat different practice is favored by Todd (**297**), who advocates training children in a "graphic vocabulary" consisting of stereotyped symbols, forms, and figures. Like letters of an alphabet, these forms may be combined in an almost infinite number of different ways in borders, posters, and other forms of decorative design. Powel (**249**) and her associates at the Cleveland Museum of Art, Voss (**305**), and many others stress that acquaintance with the work of great masters as well as instruction in art appreciation has a favorable effect on the art work of children.

Mott (**214**) with children, and Geck (**100**) with college students have shown that adding kinesthetic experience to visual and auditory impressions improves the quality of drawings. Mott had children exercise members of the body as a group "game" before drawing the human figure. Geck emphasized specific tactual as well as kinesthetic experience by having students explore manually a modelled human head before sketching it.

Dubin (**70**), whose experimental and control subjects were carefully paired with respect to age, sex, and intelligence, found that training definitely planned to help children advance from an earlier to a later stage of development in drawing (as these stages have been outlined by psychologists) was more effective than other methods commonly used in schools. Okerbloom (**235**) has presented some data on the effectiveness of Hoyt Sherman's method of teaching art. The essential feature of the method consists in training students to perceive a scene or an object as a whole in which some feature is taken to be predominant with other parts subordinated to the central theme, but nevertheless dynamically related to it.

DRAWING AS A "PROJECTIVE TECHNIQUE." I. THEORY AND METHODS

Many present-day psychologists believe that drawings and paintings as well as play and other forms of spontaneous behavior reveal children's feelings and desires. Such activity, they hold, expresses not only the needs and emotions dominant at the time, but also the more deep-seated and lasting characteristics known as "personality." Although many earnest attempts to validate this hypothesis have been made, the statistical methods used have frequently been ill-chosen. Moreover, adult associations and concepts have in their turn been projected into the interpretations of the children's "projections" in many instances. This fact provides additional evidence as to the general soundness of projective theory, but it does not induce confidence in the accuracy of the interpretations! In far too many instances, the reasoning has been circular as, for example, when the same group of cases is used to validate a hypothesis as was originally used for formulating it. Nevertheless, flimsy though many of the studies are, the accumulated evidence suggests that the basic concept is sound. Useful and valid signs by means of which dependable conclusions may be reached are, for the most part, still to be ascertained. Many have been reported (Bell [25]), but there is little uniformity with respect to their character and reputed meaning. Some agreement in interpretations is to be expected as a result of the common acceptance of certain classes of symbols by child psychologists who have been strongly imbued with psychoanalytic theory. That smearing of colors should be accepted as a sign of anal eroticism, that a preference for red should be thought to indicate feelings of aggression, or that excessive use of dark brown or black should denote depression are assumptions easily understood in terms of common adult associations. As far as children are concerned, however, little evidence is available beyond that of selected case histories. The drawings made by these children may or may not be representative of those made by others whose behavior is similar. Drawings of a like type may be equally common among children who show no indications of the behavior such drawings have been assumed to symbolize. All this has been said many times before, but so much of the work with projective methods has ignored the caution that it seems well to repeat it here.

Of the many studies that have dealt with the art of nursery school children, by far the most elaborate in its conception and presentation is the handsome two-volume report by Alschuler and Hattwick (4). Volume II presents the statistics on which the conclusions of Volume I presumably are based, together with more complete biographical data for each child. Here there is evidence of considerable confusion between that which may be true for an individual and that which is characteristic of the group to which he belongs. An example is a case described in Volume I. A little girl designated as Aileen, who was shuttled back and forth between the homes of her divorced parents with occasional side trips to the home of a grandparent or other relative, repeatedly painted an ovular red mass which she said was a "house." Actually, 133 of the 187 paintings made during her first year in nursery school were of this type. The authors state, without giving figures, that these paintings were more likely to be made during periods when the child was unusually depressed and upset, and that in subsequent years when her paintings had advanced to the pictorial stage, she not infrequently reverted to the old pattern at times when her home life was more than usually unhappy. One would have liked some figures here, yet the following explanation (Vol. I, p. 12) is at least plausible:

> It is not difficult to understand Aileen's unhappiness and her constant anxiety about "home". . . . In her persistent painting of the ovular red mass, as the full data substantiate, Aileen was expressing her need and deep craving for a home which would give her adequate protection and love.

The present writers are inclined to think that in view of the very large amount of material available for each one of the 150 children, the authors might have done better to have used statistical methods which have been worked out for the study of the individual case, together with some interchild comparisons, without attempting to pool the findings in the manner described in Volume II. The latter method may, in fact, be misleading. There is no good reason to suppose that the symbolic language of children is universal; it may well differ from one child to another. If this is true, then the only generalizations possible are in terms of principles, not of specific symbols. The criticisms of Harris (**118**), Anderson (**15**), and others with respect to the discrepancies between the interpretative material and the quantitative evidence of this study are assuredly warranted. Nevertheless, the material gathered should be of much value if the sympathetic insight of the authors were supplemented by a greater degree of statistical acumen.

Appel (**16**) describes an interesting method of utilizing drawing as an aid to securing free verbal expression from a child in the course of a psychiatric interview. The procedure, which is presented in detail, involves alternate drawing by child and psychiatrist with gradual approach to the more intimate details of home life about which children are often unwilling to speak freely. England (**87**) has studied parental identification in children by means of childlike stick figures. Schliebe (**267**) was interested in determining how children depict states of feeling in their drawings. Six drawings were secured from each of 478 children between the ages of four and eighteen years. First they were told to draw a tree. This was used as the standard with which other drawings were to be compared. The children were then told to draw a cold tree, a happy tree, a frightened tree, a sad tree, and a dying tree. Comparisons were made in terms of height and width, the direction in which the branches typically pointed, and a number of other features. The children were then asked their reasons for making the kind of drawings they did. Recall or imagination of their own kinesthetic sensations under similar conditions was the typical reason given.

The child's concept of the self is commanding more attention in the literature of recent years. The significance of drawing for the self-concept has been discussed by Dallinger (**61**) and by Wolff (**321**). Wolff, like Appel, suggests that the child's drawing of his family may show how he feels about his psychological status in that group. Machover's (**189**) fairly elaborate theory is based on the hypothesis that the self is projected into drawings of the human figure and that interpretation can be based on direct analogy. Brown and Goitein (**43**) believe that a subject's self portrait, drawn while blindfolded, will reveal in the distortions certain characteristic personality "needs." Spielrein (**280**), Wolff (**319**), and Mira (**206**) have also noted the clinical significance of drawings made without visual inspection of the results.

Bühler (**47**) has used the familiar Ball and Field Test from the Stanford-Binet to study the effect of emotional disturbances. Wechsler and Hartogs (**311**) have used mirror drawing for a similar purpose. The Bender-Gestalt Test (**28**) already mentioned has been widely used clinically, but has not been much discussed in the literature, although Billingslea (**36**) has reported an attempt at "standardization." Harms (**116**) has proposed three methods which he considers likely to prove useful in the study of neurotic children: (1) single lines drawn to denote words having affective value (a method later reported more fully), (2) drawing a picture to illustrate any word of his choice from a list of strongly affective terms, and (3) the "face" test, in which the child is asked to make one picture illustrating the thing he likes best and another illustrating the thing he most strongly dislikes. These devices are not offered as finished products but as suggestions that have appeared promising in a number of preliminary cases of which examples are given. England (**86**) asked a group of ten- to fourteen-year-old children to "draw the most important event in your life." Twenty-seven per cent of all the drawings were of fears, and the great majority of drawings of fear situations reflected traumatic episodes. On the basis of this, England is inclined to the notion that fears among older children are preponderantly concrete and based on trauma. Most child psychology textbook writers tend to hold that fears in older children are principally imaginative and based on associative learning. He does not question the difficulty or ease with which different kinds of complex events could be depicted. Williams (**315**), who was also interested in questions of methodology, found that drawings made by maladjusted children when alone in a room without distracting influences are more revealing than those made in the presence of the examiner. However, this finding should be checked; it may vary with the interpersonal relations of examiner and child. Most clinical psychologists hold that the drawing behavior itself must be observed.

Zilahi-Beke (**328**) made use of the works of Michelangelo to show how childhood experiences may influence the character of the art work of an adult. He explains the gigantic size of the figure of the young David (with whom the artist presumably identifies himself) on the basis of Michelangelo's jealousy and resentment of the favored position of his older brother. Shock and grief at the loss of his mother at the age of six is suggested as an explanation for the loving care with which he depicted the figure of the Virgin in the Pieta group as well as for her very youthful appearance, which would correspond to his memories of his own mother at the time of her death. These speculations are interesting and, to some extent, plausible, but there is, of course, no way to verify them.

On the basis of a critical study of the drawings of eight children, four of whom were popular with their playmates and four unpopular, Elkisch (**80**) developed a method of scoring by which a child's social qualities might be diagnosed. Following are the criteria used: rhythm vs. rule; complexity vs. simplexity; expansion vs. compression; and integration vs. distintegration. Examples and definitions of each type are given. Until the scale has been given more extensive trial, little can be said about its usefulness, particularly since no outside validation was attempted by the author. In a second article (**81**), Elkisch reiterates her position that a single drawing is insufficient evidence for sound diagnosis and that form rather than content should be the basis for judgment. This latter view, that formal elements are more significant for personality than meanings *per se*, is reminiscent of the Rorschach. Unlike many others of the psychoanalytically inclined who favor "meanings," Schilder and Levine (**266**) apply Freudian symbolism to the formal elements of line, color, angles, borders, contours, balance, and the like, rather than to interpretations of content.

A similar emphasis on formal "scoring elements" appears in Waehner's study (**308**), in which students were asked to draw human faces and figures of specified types. Her scoring sheet suggests the Rorschach even more strongly than the scheme of Elkisch. The statistical treatment is crude but as far as the data go, they suggest that the method has some value in differentiating personality types. Munroe, Levinson, and Waehner (**216**) with college students as subjects found that the Waehner system of interpreting drawings predicted standing on the Rorschach with a degree of accuracy that exceeded the one per cent level of confidence (chi-square method). Pfister (**244**) used color and movement in crayon drawings as diagnostic elements.

Buck offers the H-T-P Test (**45, 46**), which requires the subject to draw a picture of a house, a tree, and a person after which he responds to questions concerning the factors determining the character of his drawings. The method is discussed by Landisburg (**152**), together with an account of six cases and samples of their drawings. Buck's monograph (**46**) supplies detailed definitions of scoring categories and rules for scoring. Thus, a basis for quantitative study is afforded. Unfortunately, Buck's own statistical criteria for denoting certain characteristics as unusual, while consistently applied, appear to have no basis in statistical logic (e.g., satisfying a particular level of significance). In Buck's monograph an appendix by Payne makes some suggestions concerning the use of color.

A more abstract approach to the problem is described by Harms (**117**). Approximately ten thousand subjects ranging in age from the kindergarten to the adult level were asked to draw lines representing certain words, chiefly verbs and adjectives such as "walking," "cry," "silent," etc. His immediate task was to reduce the material to a series of objectively defined types which could be used as a basis for further work. Six forms were eventually delimited: (1) *monographical*, (2) *cursive*, in which the same figure was repeated several times in a running line, (3) *pictographical*, (4) *script*, in which a long line was drawn with undulations or circles such as to suggest writing, (5) *spatial*, in which the subject attempted to convey the quality by a change in the direction or form of the line, and (6) *final*, in which the only guiding principle seemed to be that of bringing the line to an end. Whether or not the method may prove useful in the actual differentiation of personality types remains to be seen, but the study is one of the most carefully worked out in the field. Using a very similar method, Krauss (**147**) found that when lines drawn by one set of students to depict various emotional states were shown to a second group, correct identification was made in 72.5 per cent of the cases.

Perhaps the most daring of the methods proposed for the analysis of child personality by means of drawing is that given by Wolff (**321**) in the Table of Graphic Elements on pages 223–226 of his *Personality of the Preschool Child*. Although the table is preceded by a few precautionary remarks, the list is presented as if it were wholly factual. Unfortunately, it is likely that these cautions will be unheeded by the enthusiasts and the scientifically unsophisticated who may be impressed by the plausibility of the argument. Some seventy characteristics of children's drawings are arranged in columns with the personality attribute supposedly indicated thereby placed in a parallel column. Examples are:

Preference for circular movements—Balance, changing moods, evading any decision, manic-depressive
Preference for shadings—Tactile sensitivity
Interruptions in strokes—Stubbornness, negativism

It is to be noted that most of the drawing characteristics named represent continuous rather than discrete variables, though they are not so indicated here. Apart from descriptive accounts of individual cases, objective evidence for the validity of these relationships is wholly wanting. Wolff promises a detailed description of the scoring and analytic method, together with research evidence in a later volume on drawing and handwriting (**322**). From our viewpoint, it is regrettable that this later volume, although presenting some ingenious hypotheses on the projective character of handwriting, contains only a brief mention of drawing. Actually, it refers back to the earlier text for the detailed discussion and data! This illustration is presented to indicate the all too frequent state of affairs in the work with projective techniques. Ingenious hunches and anecdotal "evidence" appear, but not painstaking objective analysis conducted with sufficient cases, or repeated samples.

Napoli (**219**) has been rather more cautious in his conclusions. He describes points to be observed which he thinks may prove useful for diagnostic purposes, but does not ascribe hard and fast meanings to them.

Brick's analysis (**40**) of the significance of color preferences shown in the drawings and paintings of children seems likewise to be more dependent upon intuition than upon facts. The same can be said of Machover's interpretation (**189**) of the human figure drawings of her subjects. Further discussions of the projective theory as applied to child art have been given by Tomlinson (**299**), Bruen (**44**), Kienzle (**142**), and others. A brief note by Green (**110**) on the response of a boy of three years to the first question known to have been put to him concerning the meaning of his scribbles on his toy blackboard holds some theoretical interest. Pointing to the drawing, his father asked, "What is that?" The child hesitated a moment and then replied, "It's a thinking."

Upon the assumption that in the art products of the blind, from whom the visual appearance of the external world is largely or wholly hidden, some insight into the nature of creative experience may be had, Lowenfeld (**172**) studied a large number of the drawings, paintings, and clay models made by blind and partially sighted children and adults. He questioned them about their mental processes while drawing, and inquired into the reasons determining their manner of drawing. From his studies he concludes that there are two sources of experience, leading to two types of creative expression. The first has its origin in visual or other types of perception and in the aesthetic pleasure directly experienced thereby; the artist attempts to reproduce the object as it appears and thus convey a similar pleasure to others. This type of art work Lowenfeld designates as the optic or visual type. The other type of experience is personal. It consists of the feelings and emotions aroused by the total situation; the artist primarily attempts to convey these feelings in his work. In Lowenfeld's terminology, such products are classed as "haptic." They differ from the type previously described in their lessened fidelity to objective reality and their greater fidelity to some definite core of meaning. This meaning is frequently, though not always, discernible to an outsider. There is a centralization of attention upon those features that convey the meaning, with corresponding neglect of other parts that are equally open to visual perception. Lowenfeld's discussion of these factors and his exposition of them in the many remarkably fine color prints contained in the Appendix is highly stimulating and should be read by all who are interested either in projective theory or in the psychology of art. Wiggin (**313**) in a paper not yet published has developed a combination of a test based on drawings and a rating scale, which together are designed to distinguish between the two types of graphic expression. Quantitative data on this test are insufficient to determine its usefulness. In a more recent book (**174**), which is designed as a textbook for teachers, Lowenfeld has placed more emphasis upon the developmental aspects of drawing together with notes on the possible significance of color, line, and form at the various stages. Like the preceding book, the volume is profusely illustrated.

Thus, the evidence accumulates that children's drawings provide significant cues for personality diagnosis. Thus far, however, no completely developed characterological system based on this evidence has appeared. Wartegg (**309, 310**) offers the only characterology based on drawings which has come to our attention. Like many of the German characterological systems developed during the thirties, it is politically oriented. Alschuler and Hattwick, and Wolff come closest to it in this country, though the thinking of Buck and of Machover have possibilities of enlargement into complete theories, in which a psychology of drawing is integrated with a psychology of personality. All this material needs more extensive validation.

DRAWING AS A "PROJECTIVE TECHNIQUE." II. CLINICAL USE WITH MALADJUSTED CHILDREN AND ADOLESCENTS

Many have expressed the belief that disturbed or maladjusted children may reveal signs in their drawings from which the nature and causes of their difficulties can be inferred. Zachry (**325**), Levy (**163**), Sigg-Boeddinghaus (**276**), Oberlin (**234**), and Dunnett (**73**) may be particularly mentioned for their clear discussions of the principles involved, although they present little or no new material.

Questions of methodology have interested a number of persons. For the most part, their material has been discussed in the preceding section except for a small number of papers which deal specifically with the question of diagnosing maladjustment. Drawing as a therapeutic treatment for maladjusted children has been discussed in several studies. In general, the theory assumes that the drawing serves as a form of mental catharsis whereby the child works out his difficulties. The leading exponent of this point of view is Naumburg (**223**, **224**, **225**, **226**, **227**, **228**, **229**), who has set forth her theories very lucidly and illustrated them by means of a series of detailed case studies. Several of these articles were later republished in monograph form (**230**). In each of the case studies, she has described in detail the stages in drawing by means of which the disturbed children exhibited their progress toward recovery. These stages differed from child to child according to his level of maturity and the nature of his problem but, in spite of occasional regressions to an earlier stage, Naumburg believes that progress can be observed readily by one who is skilled in reading the signs. A single example will suffice. A nine-year-old boy showing many indications of nervous and emotional disturbances on first entering the clinic was interested mainly in copying landscapes of a relatively peaceful character. Later, as his repressed feelings were brought to light, he began to prefer scenes showing volcanos in eruption, storms at sea, and the like. He then began to make original drawings of ghosts, devils, and skeletons, which were drawn with much attention to detail. At this stage, his drawing was frequently accompanied by profane or obscene remarks. He then began to draw fire scenes. At this time, his hostility toward his foster parents was openly expressed. Thereafter, he began to delight in drawing circus themes and roles in which his improved ego development could readily be observed. Naumburg states that at all times, his drawings and the remarks which accompanied them showed a clear distinction between the real and the imaginary. This was an important point in eliminating the diagnosis of schizophrenia. Bender (**27**) stresses the value of art work as a diagnostic and therapeutic medium for children who find it difficult to express themselves verbally. McIntosh and Pickford (**188**) have given a fairly detailed report of a single case in which drawing was found to have a definite therapeutic value. Less complete information about several other cases is included. Reports of single cases have also been given by Merzbach (**204**) and by Edelston (**77**).

The use of drawing as a means of establishing rapport between child and psychiatrist has been noted by several writers, among them Freeman (**98**) and Appel, whose work has already been mentioned. Arlow and Kadis (**18**) take the same view with respect to finger painting. In all of the younger children described by Witmer and her associates (**318**), drawing was found to be an important means to that end. Its use for diagnostic and therapeutic purposes is also noted in a number of the cases.

McIntosh (**187**) takes the position that drawing may be looked upon as a form of psychoanalysis and presents a number of interesting case reports in support of her thesis. Bender and Rapoport (**30**) divided the animal drawings of children into two broad categories, aggressive and non-aggressive, with several sub-categories under each head. Their zeal for finding relationships was hampered by the small number of cases studied (only seventy-five in all), in view of which their conclusions are questionable, though the procedure is suggestive. Phillips and Stromberg (245) studied the finger-paintings of twenty-five carefully matched pairs of subjects. One member of each pair was taken from a detention home for delinquents; the other from a nearby high school. Paintings were scored on the basis of predominating color, area of painting, choice of subject, and a number of similar characteristics. Differences between the delinquents and the controls reaching or exceeding the 5 per cent level of confidence were found for fourteen of the seventeen categories considered. The authors are properly cautious in drawing conclusions from these findings, for, as they point out, factors not considered in the matching may have been chiefly responsible for the differences found.

Lembke (161) compared the drawings of seventeen pairs of children selected by their teachers as being the boldest and the shyest in their classes. Although the small number of cases render generalizations hazardous, it is interesting to note that the differences which appeared run contrary to what might be expected on the basis of surface reasoning. The shy children used brighter colors than the bold, and showed a definite preference for complementary colors. The bold children used darker colors—in several cases a combination of violet and dark brown. In another small group of cases, Traube (300) found the use of violet and brown to be indicative of "depression." Alschuler and Hattwick (4) note a developmental shift in color preference, younger children preferring warmer colors (red, yellow), older children preferring cooler colors (blue, green). They maintain that among children of a given age, the less well adjusted and more immature will prefer the warmer colors, and the more over-controlled, "grown-up," well-behaved children will use the cooler colors. Ruesch and Finesinger (260) related the Rorschach color response to the use of color in paintings. The feeble-minded cases used more red in the drawings than normals, but gave no red differential in Rorschach responses. Introverts preferred gray in painting. Obviously, there is little agreement among these findings. The significance of color preference has long been a matter for conjecture and belief; its diagnostic value needs more complete verification.

England's (85) comparison of the drawings of "middle class" normal, mentally retarded, neglected, and delinquent children seems to have yielded few differential results, a fact that may perhaps be ascribed to his failure to control even such obvious matters as age and sex.

Vogelsang (304) classified the art products of boys in a correctional institution, according to the themes depicted. "Erotic" drawings made up 43 per cent of the total; adventure, 9.6 per cent; caricatures, 8 per cent; autobiographic fantasy, 6.6 per cent; figurative drawings including animals, landscapes, abstract designs, etc., 32.6 per cent. The longer had been the period of incarceration, the greater the number of art products made—a fact which suggests that such work, like the daydream, afforded a means of escape from the accumulating unpleasantness of their institutional life.

DRAWING AS A "PROJECTIVE TECHNIQUE." III. THE ART PRODUCTS OF PSYCHOTIC CHILDREN AND ADOLESCENTS

Inasmuch as true psychotic states are rare before the latter part of the adolescent period, the literature on the drawings of psychotic children is scanty. The four reviews by Anastasi and Foley on the artistic behavior of the abnormal (**11**, **12**, **13**, **14**), cover the field up to about 1940, but, as would be expected, their reports deal mainly with the work of adults. We have checked their bibliographies against ours and have eliminated all duplications except those dealing with the art of children and a small number of others in which significant comments or comparisons with child art are made. We have not attempted to cover the fairly large body of literature on the art of abnormal adults since the Anastasi and Foley reviews except in a few instances of the type just mentioned. For example, we have listed Fleming's article (**92**) on the use of finger painting in mental hospitals because of the popularity of the method as a projective device for use with children. Mohr (**208**) has stressed the significance of the strange and unusual features so often found in the drawings of the mentally ill and has pointed out that while patients rarely depict their hallucinations overtly, they do so covertly by means of these peculiar symbols that are incomprehensible to anyone but themselves. This is as true of children as of adults.

Occasionally the drawings of a child resemble closely those of a schizophrenic patient. Cameron (**51**) ascribes this resemblance to the failure of the usual mechanisms whereby the drawing serves a form of social communication. It thus takes on forms that are peculiar to the individual as is true in the cases of those made by the schizophrenic.

Fingert, Kagan, and Schilder (**91**) used the Draw-a-Man Test as a means of measuring improvement in schizophrenic patients following insulin or metrazol shock. Although forty cases were studied, only the detailed results for several individual subjects said to be typical are reported. A steady improvement in score during the period of recovery from shock is apparent in all these cases. It is unfortunate that no control drawing seems to have been secured previous to the administration of the shock. Schilder and Levine (**266**) trace the relationship between drawing and psychic disturbance to the basic motor drives of the individual which, in the seriously disturbed, modify his handling of the abstract principles of drawing—curves, angles, straight lines, balance, and the like.

Abel (**1**) asked several groups of institutionalized patients to produce a design within a 4 in. by 6 in. rectangle by the use of exactly 19 straight and 6 curved lines. She found that schizophrenics were particularly concerned about getting in the exact number of lines, but were typically unable to make a balanced design; their drawings were unorganized. This lack of organization was especially marked in paranoids and to a less degree in subjects classed as "normal, showing paranoid trends."

There are a few published case studies of the drawing behavior of individual children and adolescents tentatively diagnosed as psychopaths or at least as extreme neurotics. Naumburg (**223**) has given an account of the drawings of an adolescent girl suffering from anxiety hysteria with amnesia. The factors giving rise to her complaint were revealed in her drawings before she was able to express them verbally. Merzbach (**204**) has described the work of a thirteen-year-old boy during a temporary psychotic state in which he felt impelled to express himself by drawing. Analysis and discussion of the drawings showed the boy to be suffering from intense anxiety, but, Merzbach felt, definitely exclude the diagnosis of schizophrenia. It is quite possible that some of the unusual cases described in other reports were in reality exhibiting the early stages of a psychosis not yet diagnosed as such.

Bender's report on the use of the Goodenough test among children suffering from chronic encephalitis has already been mentioned. This test was also used by Berrien (**33**) in a comparison of the drawings of three groups of abnormal children including 24 post-encephalitics, 13 diagnosed as having "psychopathic personality," and 15 borderline mental defectives. The Binet test, apparently, was not given. The post-encephalitic cases showed a number of the characteristics noted by Goodenough as possible indicators of neurotic or psychotic states; most notably reversal of sex characteristics and combinations of mature and primitive characteristics in the same drawing. An item analysis was also attempted, but the number of cases was too few and the spread of ages too great to lend much value to this part of the study. Chase (**56**) found that among the schizophrenic groups she studied, the hebephrenics ranked lower than the paranoids on the Goodenough test and that both were inferior to normal subjects of similar age and education.

It may be noted that most of the reports on the art of the abnormal have to do with the schizophrenic group. However, particularly in the case of children, studies of those suffering from some form of brain damage caused either by accident or disease are coming to occupy a fairly important place in the literature (see p. 377, this paper). The reason for the large number of reports based upon schizophrenic patients is two-fold. In the first place, as Kris (**149**) and others have shown, a greater number of these patients are interested in drawing or other forms of artistic activity. In the second place, their drawings seem to be more distinctive in type, more clearly set off from those of normals, than are those of other clinical groups. For many schizophrenics, art becomes a means of covert self-expression. By means of seemingly inappropriate and senseless symbols, the meanings of which are known only to himself, the patient gives vent to his troubles, communicates his feelings to the insensate paper, and in so doing finds a measure of relief. From the standpoint of this survey, too, the drawings of the schizophrenic are of special interest, since this is the only form of functional mental disease that occurs in numbers great enough to be appreciable during the years of later childhood and adolescence.

More dynamic theories of personality in recent years have reinstated to some extent the belief that personality forces may undergo considerable realignment as a person departs more and more from his usual habits of perception and judgment, that somewhere along the behavioral continuum from the normal to the abnormal a fairly sudden reorganization of the personality occurs. One gets this impression from a consideration of the art of seriously disturbed or psychotic subjects. Bizarre features appear; the total effect of pictures by such patients is quite different. The process by which a severe neurosis or a psychosis develops has not, to our knowledge, been studied in a suitable collection of drawings. But the process, as this review points out, has been studied in reverse. As the patient improves, rather sudden and dramatic shifts in content and technique of drawings have been observed. This possibility leads us to question whether interpretive principles suitable for study of the art work of clinic patients can be applied directly in the personality analysis of well-adjusted children. Brown and Goitein (**43**), for example, conclude from studies of psychotic subjects that when a normal person's drawing fits a particular abnormal category, his personality will show trends intermediate between the normal and the not-normal which fit that psychiatric group.

Even though the continuity exists, the variance in "normal" expressions of abnormal deviations may be so slight as to defy our present methods of observation. It may be that "signs" of behavior as contrasted with samples of skills[10] are best considered as attributes rather than as variables, or that they must be judged more ideographically than nomothetically. Such possibilities, at least, may indicate why attempts to validate observations in the clinic by discovering an analogous though smaller variance among normal subjects have proved disappointing.

CONCLUSIONS

In the course of the twenty years since the last report on this subject appeared in the *Bulletin*, interest in the psychological aspects of children's drawings has continued to be strong but its direction has been somewhat modified. Chiefly, this change may be attributed to the recent upsurge of interest in projective theory and method. Although both the psychological use of the term "projection" and its underlying concepts had been well known since the early writings of Freud, and although a number of well-known "projective" tests were developed previous to 1939, the recognition that methods of this type are based on a different psychological theory from those of the more conventional type, waited for Frank's definition in 1939.[11] As is not infrequently the case, the immediate result of classifying and naming was a marked upswing of interest in these procedures. Although reference to the bibliography will show that the practical use of children's drawings as a projective device began long before Frank's article was written, the underlying theory upon which these earlier studies were based was rarely stated explicitly.

---

[10] Goodenough, Florence L. The appraisal of child personality. *Psychol. Rev.*, 1949, **56**, 123–131.

Other changes in the types of studies that have appeared since the publication of the earlier article may be noted. Interest in the developmental aspects of children's drawing has shifted from the broad outline to more specific details of form and line, composition, design, and the use of various media, such as clay or finger painting. There has been far more concern than formerly over questions of methodology, particularly among those whose primary interest is in the field of tests and measurements. It is unfortunate that this concern has not been shared by the majority of those interested in projective methods.

The possibility of utilizing children's drawings as aids to the solution of problems in general psychology has also been more clearly realized. For obvious reasons, visual perception and visual memory have been the areas most frequently studied by drawings, but the ramifications of these topics are many. The relationship between perception and the characteristics of the perceiving individual is a fascinating topic, of which only exploratory studies have been made. Finally, the large body of literature on the art of the abnormal leads directly to the art of childhood, for the psychotic adult was once a child who, in all likelihood, drew pictures. Whether or not these pictures differed in any determinable way from those drawn by children who were destined to lead more normal lives is unknown, but the possibility is worth investigating. Actual drawings, if preserved, are free from the halo effect that is likely to be found in verbal reports by persons whose memories of the patient's early behavior may be influenced by a knowledge of his later difficulties.

---

[11] Frank, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, **8**, 339–418.

Study of the art of children is now well past the stage at which a useful purpose can be served by simple tabulations of subjects drawn by miscellaneous groups of children, or the reiterations of developmental sequences that were worked out a quarter of a century ago. The urgent need at the present time is a series of objectively defined and reliably measurable categories which may be used for the testing of hypotheses and the determination of relationships. The existence of such measures would by no means free the research worker or the clinician from a need for the type of insight that can be gained only from intelligent experience. Such categories, however, would enable him to design his experiments in such manner that the validity of his claims could be more readily judged by others. A variety of measures suitable for many different purposes in the fields of experimental and clinical psychology would be desirable, for the art of children and youth may be considered from many different aspects, and drawings vary widely from one experimental set-up to another. The field is wide and the way is open.

BIBLIOGRAPHY

1. Abel, Theodora M. Free designs of limited scope as a personality index. *Character & Pers.*, 1938, **7**, 50–62.
2. Abel, Theodora M., & Sill, Jane B. The perceiving and thinking of normal and subnormal adolescents and children on a simple drawing task. *J. genet. Psychol.*, 1939, **54**, 391–402.
3. Alschuler, Rose H., & Hattwick, La Berta W. Easel painting as an index of personality in preschool children. *Amer. J. Orthopsychiat.*, 1943, **13**, 616–625.
4. Alschuler, Rose H., & Hattwick, La Berta W. *Painting and personality, a study of young children.* Chicago: Univ. of Chicago Press, 1947. Two vol.
5. Alschuler, Rose H., & Hattwick, La Berta W. Understanding children through their paintings. *Understanding the Child*, 1947, **16**, 98–101.
6. Ames, Louise B. The Gesell Incomplete Man Test as a differential indicator of average and superior behavior in preschool children. *J. genet. Psychol.*, 1943, **62**, 217–274.
7. Ames, Louise B. Free drawing and completion drawing: a comparative study of preschool children. *J. genet. Psychol.*, 1945, **66**, 161–165.
8. Anastasi, Anne, & Foley, J. P., Jr. An analysis of spontaneous drawings by children in different cultures. *J. appl. Psychol.*, 1936, **20**, 689–726.
9. Anastasi, Anne, & Foley, J. P., Jr. A study of animal drawings by Indian children of the North Pacific Coast. *J. soc. Psychol.*, 1938, 9, 363–374.
10. Anastasi, Anne, & Foley, J. P., Jr. The study of "populistic painters" as an approach to the psychology of art. *J. soc. Psychol*, 1940, **11**, 353–368.
11. Anastasi, Anne, & Foley, J. P., Jr.

A survey of the literature on artistic behavior in the abnormal. I. Historical and theoretical background. *J. gen. Psychol.*, 1941, **25**, 111–142.

12. Anastasi, Anne, & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal. II. Approaches and interrelationships. *Ann. N. Y. Acad. Sci.*, 1941, **42**, 1–111.
13. Anastasi, Anne, & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal. III. Spontaneous productions. *Psychol. Monogr.*, 1940, **52**, No. 6, 71 pp.
14. Anastasi, Anne, & Foley, J. P., Jr. A survey of the literature on artistic behavior in the abnormal. IV. Experimental investigations. *J. gen. Psychol.*, 1941, **25**, 187–237.
15. Anderson, J. E. Personality organization in children. *Amer. Psychologist*, 1948, **3**, 409–416.
16. Appel, K. E. Drawings by children as aids to personality studies. *Amer. J. Orthopsychiat.*, 1931, **1**, 129–144.
17. Appia, B. La représentation humaine dans les dessins d'enfants noires. (The human being in drawings by Negro children.) *Bull. Inst. franç. Afrique noir*, 1939, **1**. (*Psychol. Abstr.*, 1946, **20**, No. 3341.)
18. Arlow, J. A., & Kadis, Asja. Finger painting in the psychotherapy of children. *Amer. J. Orthopsychiat.*, 1946, **16**, 134–146.
19. Barnhart, E. N. Developmental stages in compositional construction in children's drawings. *J. exp. Educ.*, 1942, **11**, 156–184.
20. Baumgarten, Franziska. Supranormales Zeichnen eines Kindes. (Supernormal drawing of a child.) *Z. Kinderpsychiat.*, 1936, **2**, 182–189.
21. Baumgarten, Franziska, & Tramer, M. Kinderzeichnungen in vergleichend psychologischer Beleuchtung. (Child drawings in the light of comparative psychological study.) *Z. Kinderpsychiat.*, 1943, **9**, 161–220.
22. Beach, Vera, & Bressler, Mary. Organizing a painting set-up for young children. *J. exp. Educ.*, 1944, **13**, 77–80.
23. Beck, O. W. *Self-development in drawing as interpreted by the genius of Romano Dazzi and other children.* New York, London: G. P. Putnam's Sons, 1928.
24. Belart, W. Die sinnespsychologischen Grundlagen seiten- und hohenverkehrter Kinderzeichnungen. (The sensory basis of lateral and vertical inversions in children's drawings.) *Z. Kinderpsychiat.*, 1943, **10**, 8–13.
25. Bell, J. E. *Projective techniques; a dynamic approach to the study of personality.* New York: Longmans, Green and Co., 1948.
26. Bender, Lauretta. Gestalt principles in the sidewalk drawings and games of children. *J. genet. Psychol.*, 1932, **41**, 192–210.
27. Bender, Lauretta. Art and therapy in the mental disturbances of children. *J. nerv. ment. Dis.*, 1937, **86**, 249–263.
28. Bender, Lauretta. *A visual motor gestalt test and its clinical use.* Research Monographs, No. 3. New York: American Orthopsychiatric Assn., 1938.
29. Bender, Lauretta. The Goodenough test (drawing a man) in chronic encephalitis in children. *J. nerv. ment. Dis.*, 1940, **91**, 277–286.
30. Bender, Lauretta, & Rapoport, J. Animal drawings of children. *Amer. J. Orthopsychiat.*, 1944, **14**, 521–527.
31. Bender, Lauretta, & Wolfson,

W. Q. The nautical theme in the art and fantasy of children. *Amer. J. Orthopsychiat.*, 1943, **13**, 462–467.

32. Berdie, Ralph F. Measurement of adult intelligence by drawings. *J. clin. Psychol.*, 1945, **1**, 288–295.

33. Berrien, F. K. A study of the drawings of abnormal children. *J. educ. Psychol.*, 1935, **26**, 143–150.

34. Biber, Barbara. *Children's drawings: from lines to pictures.* New York, Bureau of Educational Experiments, Cooperating School Pamphlet No. 6, 1934.

35. Billing, Marion L. A report of a case of inverted writing and drawing. *Child Develpm.*, 1935, **6**, 161–163.

36. Billingslea, F. Y. The Bender-Gestalt: an objective scoring method and validating data. *J. clin. Psychol.*, 1948, **4**, 1–27.

37. Bird, M. H. *A study in aesthetics.* Harvard Monographs in Education, No. 11, Cambridge: Harvard University Press, 1932.

38. Blum, Lucille H., & Dragositz, Anna. Finger painting: the developmental aspects. *Child Develpm.*, 1947, **18**, 88–105.

39. Boas, F. *Primitive art.* Cambridge, Mass.: Harvard Univ. Press, 1927.

40. Brick, Maria. Mental hygiene value of children's art work. *Amer. J. Orthopsychiat.*, 1944, **14**, 136–146.

41. Brill, M. A study of instability using the Goodenough Drawing Scale. *J. abnorm. soc. Psychol.*, 1937, **32**, 288–302.

42. Brooks, F. D. The relative accuracy of ratings assigned with and without use of drawing scales. *Sch. & Soc.*, 1928, **27**, 518–520.

43. Brown, E. A., & Goitein, P. L. The significance of body image for personality assay. *J. nerv. ment. Dis.*, 1943, **97**, 401–408.

44. Bruen, C. The artistic process. *Psyche*, 1930, **10**, 41–59. (*Psychol. Abstr.*, 1931, **5**, No. 459.)

45. Buck, J. N. The H-T-P Test. *J. clin. Psychol.*, 1948, **4**, 151–159.

46. Buck, J. N. The H-T-P technique: a qualitative and quantitative scoring technique. *J. clin. Psychol. Mon. Suppl.*, No. 5, Oct., 1948.

47. Bühler, Charlotte. The Ball and Field Test as a help in the diagnosis of emotional difficulties. *Character & Pers.*, 1938, **6**, 257–273.

48. Burkhardt, H. Über Verlagerung raümlicher Gestalten. (On the alteration of spatial form.) *Neue psychol. Stud.*, 1933, **7**, No. 3. (*Psychol. Abstr.*, 1935, **9**, No. 4394.)

49. Burton, A., & Tueller, Roma. Successive reproductions of visually perceived forms. *J. genet. Psychol.*, 1941, **58**, 71–82.

50. Cain, T. D. The objective measurement of accuracy in drawings. *Amer. J. Psychol.*, 1943, **56**, 32–53.

51. Cameron, N. Individual and social factors in the development of graphic symbolization. *J. Psychol.*, 1938, **5**, 165–184.

52. Cane, Florence. The gifted child in art. *J. educ. Sociol.*, 1936, **10**, 67–73.

53. Carroll, H. A. A preliminary report on a study of the relationship between ability in art and certain personality traits. *Sch. & Soc.*, 1932, **36**, 285–288.

54. Cattell, R. B. Personality traits associated with abilities. I. With intelligence and drawing ability. *Educ. psychol. Measmt*, 1945, **5**, 131–146.

55. Chang, K. [Application of Goodenough's intelligence test by drawing in China.] *Chung Hwa educ. Rev.*, 1935, **23**, No. 5, 49–55. (*Psychol. Abstr.*, 1936, **10**, No. 5161.)

56. Chase, Jane M. A study of the

drawings of a male figure made by schizophrenic patients and normal subjects. *Character & Pers.*, 1941, **9**, 208–217.

57. COCKRELL, D. L. Design in the paintings of young children. *Sch. Arts*, 1930, **30**, 33–39; 112–119.
58. COHEN, D. N. The Goodenough Drawing Scale applied to thirteen-year-old children. M. A. thesis, Columbia Univ., 1933. (Not seen.)
59. COOK, F. Art without sight. *Sch. Arts*, 1943, **43**, 40–45.
60. CURRAN, F. J. Art techniques for use in mental hospitals and correctional institutions. *Ment. Hyg., N. Y.*, 1939, **23**, 371–378.
61. DALLINGER, K. Über den Zusammenhang zwischen der Entwicklung des Ichbewusstseins mit dem kindlichen Zeichnen. (On the relationship between the development of the awareness of the self and the development of drawing in children.) In O. Tumitz (Ed.), *Jugendkundliche Arbeiten*, Vol. II. Langensalza: Beyer und Mann, 1928.
62. DANIELS, PARMELY C. Discrimination of compositional balance at the pre-school level. *Psychol. Monogr.*, 1933, **45**, No. 1, 1–11.
63. DATTA, A. Drawings of children. *Indian J. Psychol.*, 1935, **10**, 179–182. (*Psychol. Abstr.* 1936, **10**, No. 5559.)
64. DENNIS, W. The performance of Hopi children on the Goodenough Draw-a-Man Test. *J. comp. Psychol.*, 1942, **34**, 341–348.
65. DESPERT, J. LOUISE. Technical approaches used in the study and treatment of emotional problems in children. Part III. Drawing. *Psychiat. Quart.*, 1937, **11**, 267–295.
66. DIDE, M., & GUILHEM, M. Exploration psychologique infantile par l'expression graphique du mouvement. (A psychological investigation of graphic expression of movement by children.) *Mem. Acad. Sci., Toulouse*, 1929, **7**, 247–259. (*Psychol. Abstr.*, 1930, **4**, No. 3633.)
67. DOMS, F. P. Les échelles Gaublomme. Le graphisme et l'expression graphique. (The Gaublomme scales, writing and graphic expression.) *Sem. univ. Pédag. Univ. libre Brux.*, 1935, **1**, 171–182. (*Psychol. Abstr.*, 1936, **10**, No. 4140.)
68. DOW, MILDRED. Playground behavior differentiating artistic from non-artistic children. *Psychol. Monogr.*, 1933, **45**, No. 1, 82–94.
69. DOW, MILDRED. An observational study in a playground situation of differences in artistic personality at the child level. *Proc. Iowa Acad. Sci.*, 1933, **40**, 197.
70. DUBIN, ELISABETH R. The effect of training on the tempo of development of graphic representation in preschool children. *J. exp. Educ.*, 1946, **15**, 166–173.
71. DU BOIS, CORA. *The people of Alor.* Minneapolis: The University of Minnesota Press, 1944.
72. DUMONT, L. L'émotion esthétique chez l'enfant. (Esthetic emotion in the child.) *Bull. Instit. gén. Psychol.*, 1927, **27**, Nos. 1–3, 62–76, (*Psychol. Abstr.* 1928, **2**, No. 1066.)
73. DUNNETT, RUTH. *Art and child personality.* London: Methuen, 1948.
74. DYETT, E. G. The Goodenough Drawing Scale applied to college students. M. A. thesis, Columbia University, 1931. (Not seen.)
75. EARL, C. J. C. The human figure drawings of adult defectives. *J. ment. Sci.*, 1933, **79**, 305–328.
76. EARL, C. J. C. The human figure drawings of feeble-minded adults. *J. Psycho-Asth.*, 1933, **38**, 107–120.
77. Edelston, H. The analysis and treatment of a case of neurotic conduct disorder in a young child illustrat-

ing the value and use of drawing in child guidance technique. *J. ment. Sci.*, 1939, **85**, 522–547.

78. Eells, W. C. Mental ability of the native races of Alaska. *J. appl. Psychol.*, 1933, **17**, 417–438.
79. Eggers, M. M. Comparison of Army Alpha and Goodenough Drawings in delinquent women. M. A. thesis, Columbia Univ., 1931. (Not seen.)
80. Elkisch, Paula. Children's drawings in a projective technique. *Psychol. Monogr.*, 1945, **58**, No. 1, 1–31.
81. Elkisch, Paula. The emotional significance of children's art work. *Childh. Educ.*, 1947, **23**, 236–241.
82. Ellsworth, Frances F. Elements of form in the free paintings of twenty nursery school children. *J. gen. Psychol.*, 1939, **20**, 487–501.
83. Eng, Helga K. *The psychology of children's drawings.* New York: Harcourt, Brace, 1931. (First publishe din German, 1927.)
84. Engerth, G. Zeichenstörung bei Patienten mit Autotopagnosie. (Disorders of drawing in patients with autotopagnosia.) *Z. ges. Neurol. Psychiat.*, 1933, **143**, 381–402.
85. England, A. O. A psychological study of children's drawings: comparison of public school, retarded, institutionalized, and delinquent children's drawings. *Amer. J. Orthopsychiat.*, 1943, **13**, 525–530.
86. England, A. O. Non-structured approach to the study of children's fears. *J. clin. Psychol.*, 1946, **2**, 364–368.
87. England, A. O. Cultural milieu and parental identification. *Nerv. Child*, 1947, **6**, 301–305.
88. Fallaize, E. N. Children's drawings and primitive art. *J. Educ. (Lond.)*, 1933, **65**, 142–143.
89. Farnsworth, P. R. Ratings in music, art, and abnormality in the first four grades. *J. Psychol.*, 1938, **6**, 89–94.
90. Findley, W. G. Factor analysis of a short-item drawing test. *Psychol. Bull.*, 1936, **33**, 605.
91. Fingert, H. H., Kagan, Julia R., & Schilder, P. The Goodenough Test in insulin and metrazol treatment of schizophrenia. *J. gen. Psychol.*, 1939, **21**, 349–365.
92. Fleming, Joan. Observations on the use of finger painting in the treatment of adult patients with personality disorders. *Character & Pers.*, 1940, **8**, 301–310.
93. Flemming, E. G. Personality and artistic talent. *J. educ. Sociol.*, 1934, **8**, 27–33.
94. Foley, J. P., & Anastasi, Anne. The work of the Children's Federal Art Gallery. *Sch. & Soc.*, 1938, **48**, 859–861.
95. Fontes, V. O sobrenatural nos desenhos infantis. (The supernatural in children's drawings.) *Criança portug.*, 1946–47, **6**, 71–100. (*Psychol. Abstr.*, 1948, **22**, No. 4329.)
96. Fontes, V., Leite da Costa, M. I., & Ferreira, A. Contribuição para o estudo do teste de Fay aplicado em crianças protuguesas. (Contribution to the study of the Fay test as used with Portuguese children.) *Criança portug.*, 1944, **3**, 13–34. (*Psychol. Abstr.*, 1945, **19**, No. 1057.)
97. Fordham, M. S. C. *The life of childhood.* London: K. Paul, Trench, Trubner and Co., 1944.
98. Freeman, Mabel. Drawing as a psychotherapeutic intermedium. *J. Psycho-Asth.*, 1936, **41**, 182–187.
99. Garrison, S. C. Fine arts. *Rev. educ. Res.*, 1935, **5**, 45–47.
100. Geck, Frances J. The effectiveness of adding kinesthetic to visual and auditory perception in the teaching

of drawing. *J. educ. Res.*, 1947, **41**, 97–101.

101. Gerald, H. J. P. Inverted positions in children's drawings: report of two cases. *J. nerv. ment. Dis.*, 1928, **68**, 449–455.

102. Gesell, A., & Ames, Louise B. The development of directionality in drawing. *J. genet. Psychol.*, 1946, **68**, 45–61.

103. Gitlitz, H. B. The Goodenough Drawing Scale applied to fourteen-year-old children. M. A. thesis, Columbia Univ., 1933. (Not seen.)

104. Goodenough, Florence L. Racial differences in the intelligence of school children. *J. exp. Psychol.*, 1926, **9**, 388–397.

105. Goodenough, Florence L. Studies in the psychology of children's drawings. *Psychol. Bull.*, 1928, **25**, 272–283.

106. Goodenough, Florence L. Children's drawings. In C. Murchison (Ed.), *A Handbook of Child Psychology*, (1st Ed.). Worcester, Mass: Clark Univ. Press, 1931.

107. Graewe, H. *Untersuchung der Entwicklung des Zeichnens.* (A study of the development of drawing.) Halle: Hermann Schroedel, 1932.

108. Graewe, H. Das Tierzeichnen der Kinder. (The child's drawing of animals.) *Z. pädag. Psychol.*, 1935, **36**, 251–256; 291–300.

109. Graewe, H. Geschichtlicher Überblick der Psychologie des kindlichen Zeichnens. (Historical survey of the psychology of children's drawings.) *Arch. ges. Psychol.*, 1936, **96**, 103–220.

110. Green, G. H. A child's first attempts to interpret drawings. *J. genet. Psychol.*, 1928, **35**, 473–474.

111. Gridley, Pearl F. Graphic representation of a man by four-year-old children in nine prescribed drawing situations. *Genet. Psychol. Monogr.*, 1938, **20**, 183–350.

112. Griffiths, Ruth. *A study of imagination in early childhood and its function in mental development.* London: Kegan Paul, Trench, Trubner and Co., 1945.

113. Grippen, Velma B. A study of creative artistic imagination in children by the constant contact procedure. *Psychol. Monogr.*, 1933, **45**, No. 1, 63–81.

114. Guttmann, E., & Maclay, W. S. Clinical observations on schizophrenic drawings. *Brit. J. med. Psychol.*, 1937, **16**, 184–205.

115. Harms, E. Kinderkunst als diagnostiches Hilfsmittel bei infantilen Neurosen. (Child art as a diagnostic aid in child neuroses.) *Z. Kinderpsychiat.*, 1940, **6**, 129–143.

116. Harms, E. Child art as an aid in the diagnosis of juvenile neuroses. *Amer. J. Orthopsychiat.*, 1941, **11**, 191–210.

117. Harms, E. The psychology of formal creativeness: I. Six fundamental types of formal expression. *J. genet. Psychol.*, 1946, **69**, 97–120.

118. Harris, D. B. Review of *Painting and Personality* by Rose H. Alschuler and La Berta Hattwick. *Ment. Hyg., N. Y.*, 1948, **32**, 653–659.

119. Hartenstein, Erna. Aktuelle Kulturinhalte im Spiegel der Kinderzeichnung. (Actual manifestations of culture in the mirror of children's drawings.) *Z. Jugendk.*, 1934, **4**, 106–120.

120. Hatano, I. [The picture stories and suggestions to the education of drawing.] *Trans. Inst. Child Study* (Japanese), 1934, **16**, 811–830. (*Psychol. Abstr.*, 1935, **9**, No. 450.)

121. Haubold, M. Bildbetrachtungen durch Kinder und Jugendliche (Versuche über das Unterscheiden von Bildern verschiedenen Stiles.) (Appreciation of pictures by children and adolescents. [Experiments on the discrimination of dif-

ferent styles of pictures.]) *Neue Psychol. Stud.*, 1933, **7**, No. 2. (*Psychol. Abstr.*, 1934, **8**, No. 5301.)

122. Havighurst, R. J., Gunther, Minna K., & Pratt, Inez E. Environment and the Draw-a-Man Test: the performance of Indian children. *J. abnorm. soc. Psychol.*, 1946, **41**, 50–63.
123. Hellersberg, Elisabeth F. The Horn-Hellersberg Test and adjustment to reality. *Amer. J. Orthopsychiat.*, 1945, **15**, 690–710.
124. Hetzer, Hildegard. Die symbolische Darstellung in der frühen Kindheit. (Symbolical representation in early childhood.) *Wiener Arb. z. päd. Psychol.*, 1926, **3**, 1–92. (*Psychol. Abstr.*, 1933, **7**, No. 6025.)
125. Hicks, J. A., & Ralph, Dorothy W. The effects of practice in tracing the Porteus Diamond Maze. *Child Developm.*, 1931, **2**, 156–158.
126. Hildreth, Gertrude. Developmental sequences in name writing. *Child Develpm.*, 1936, **7**, 291–303.
127. Hildreth, Gertrude. *The child mind in evolution; a study of developmental sequences in drawing.* New York: King's Crown Press, 1941.
128. Hildreth, Gertrude. War theme in children's drawings. *Childh. Educ.*, 1943, **20**, 121–127.
129. Hildreth, Gertrude. The simplification tendency in reproducing designs. *J. genet. Psychol.*, 1944, **64**, 329–333.
130. Hinrichs, W. E. The Goodenough drawing in relation to delinquency and problem behavior. *Arch. Psychol.*, No. 175, 1935.
131. Homma, T. [The law of Prägnanz in the process of drawing figures.] *Jap. J. Psychol.*, 1937, **12**, 112–153. (*Psychol. Abstr.*, 1937, **11**, No. 4462.)
132. Hsiao, H. H. [Hsiao's revision of Goodenough's test of intelligence for children.] *Monogr. Psychol. Educ., Nat. cent. Univ.*, 1939, **4**, No. 1. (*Psychol. Abstr.*, 1941, **15**, No. 2836.)
133. Huang, I. [*The psychology of children's drawings.*] Shanghai: Commercial Press, 1938. (*Psychol. Abstr.*, 1939, **13**, No. 3889.)
134. Hurlock, Elizabeth B. The spontaneous drawings of adolescents. *J. genet. Psychol.*, 1943, **63**, 141–156.
135. Hurlock, Elizabeth B., & Thomson, J. L. Children's drawings: an experimental study of perception. *Child Developm.*, 1934, **5**, 127–138.
136. Iinuma, R., & Watanabe, K. [Observations on moving objects drawn by children. Orientation and motion of railway trains.] *Jap. J. Psychol.*, 1937, **12**, 393–408. (*Psychol. Abstr.*, 1938, **12**, No. 977.)
137. Israelite, Judith. A comparison of the difficulty of items for intellectually normal children and mental defectives on the Goodenough Drawing Test. *Amer. J. Orthopsychiat.*, 1936, **6**, 494–503.
138. Joseph, Alice, Spicer, Rosamund B., & Chesky, Jane. *The desert people.* Chicago: Univ. of Chicago Press, 1949.
139. Katô, M. [A genetic study of children's drawings of man.] *Jap. J. exp. Psychol.*, 1936, **3**, 175–185. (*Psychol. Abstr.*, 1937, **11**, No. 3000.)
140. Kerr, Madeline. Children's drawings of houses. *Brit. J. med. Psychol.*, 1937, **16**, 206–218.
141. Kerr, Rose N. Art and the young child. *Sch. Arts*, 1936, **36**, 5–8.
142. Kienzle, R. Das bildhafte Gestalten als Ausdruck per Persönlichkeit. (Pictorial forms as expressions of the personality.) Esslingen: Wilh. Langguth, 1932. (*Psychol. Abstr.*, 1933, **7**, No. 1878.)
143. Kline, L. W., & Carey, G. L. A

measuring scale for free-hand drawing: Part II. Design and composition. *Johns Hopkins Stud. Educ.*, 1933, No. 5, Part II.

144. Knauber, Alma J. A study of the art ability found in very young children. *Child Develpm.*, 1931, **2**, 66–71.

145. Kobayashi, S. [A study on a variation of facsimilies drawn by children.] *Jap. J. Psychol.*, 1937, **12**, 375–392. (*Psychol. Abstr.*, 1938, **12**, No. 979.)

146. Koga, Y., & Naka, Y. [Studies on the ability to draw.] *Kyôiku Shinri Kenkyû*, 1939, **14**, 509–519; 581–590; 749–761. (*Psychol. Abstr.*, 1940, **14**, No. 3090.)

147. Krauss, R. Über graphischen Ausdruck. Eine experimentelle Untersuchung über das Erzeugen und Ausdeuten von gegenstandfreien Linien. (Concerning graphic expression. An experimental investigation of the production and interpretation of objectless lines.) *Beih. Z. angew. Psychol.*, No. 48, 1930.

148. Kretschmann, O. Beiträge zur Frage nach der Vererbung der zeichnerischen Anlagen. (Contribution to the question of the inheritance of graphic talent.) *Arch. ges. Psychol.*, 1941, **108**, 267–316. (*Psychol. Abstr.*, 1942, **16**, No. 2645.)

149. Kris, E. Bemerkungen zur Bildnerei der Geisteskranken. (Observation on the sculptures of the mentally ill.) *Imago. Lpz.*, 1936, **22**, 339–370. (*Psychol. Abstr.*, 1936, **10**, No. 5834.)

150. Krüber, W. Über das Aufzeichnen von Formen aus dem Gedächtnis. (Concerning the drawing of forms from memory.) *Z. angew. Psychol.*, 1938, **54**, 273–327.

151. Labunskaya. Infantile drawing. *Sch. Arts*, 1934, **33**, 368–371.

152. Landisburg, Selma. A study of the H-T-P Test. *Train. Sch. Bull.*, 1947, **44**, 140–152.

153. Lange, B. Zur Psychologie der Zeichnenunterrichts. (The educational psychology of drawing.) *Int. Z. Indiv.-Psychol.*, 1933, **11**, 295–301.

154. Lark-Horovitz, Betty. Interlinkage of sensory memories in relation to training in drawing. *J. genet. Psychol.*, 1936, **49**, 69–89.

155. Lark-Horovitz, Betty. On learning abilities of children as recorded in a drawing experiment: I. Subject matter. *J. exp. Educ.*, 1941, **9**, 332–345.

156. Lark-Horovitz, Betty. Comparison of subjective and objective judgments of children's drawings. *J. exp. Educ.*, 1942, **10**, 153–163.

157. Lark-Horovitz, Betty, Barnhart, E. N., & Sills, Esther M. *Graphic work-sample diagnosis: an analytic method of estimating children's drawing ability.* Cleveland, Ohio: Cleveland Museum of Art, 1939.

158. Lassen, H. Raumauffassung und Raumdarstellung in Kinderzeichnungen. (Space perception and representation in children's drawings.) *Arch. ges. Psychol.*, 1943, **112**, 153–195. (*Psychol. Abstr.*, 1946, **20**, No. 2526.)

159. Legrün, A. Zur Deutung von Kinderkritzelein. (On the significance of children's scribbling.) *Z. Kinderforsch.*, 1938, **47**, 236–249.

160. Leighton, Dorothea, & Kluckhohn, Clyde. *Children of the people.* Cambridge, Mass.: Harvard Univ. Press, 1947.

161. Lembke, W. Über Zeichnungen von "frechen" und "schüchternen" Schulkindern. (The significance of drawings by "bold" and "shy" pupils.) *Z. päd. Psychol.*, 1930, **31**, 459–462.

162. Lerner, E., Murphy, Lois B., *et. al.* Methods for the study of person-

ality in young children. *Monogr. Soc. Res. Child Develpm.*, 1941, **6**, No. 4. Washington, D.C.: National Research Council.

163. Levy, J. The use of art techniques in treatment of children's behavior problems. *J. Psycho-Asth.*, 1934, **39**, 258–260.

164. Levy, L. R. The function of the Goodenough Drawing Scale in the study of high school freshmen. M. A. thesis, Columbia Univ., 1931. (Not seen.)

165. Lewerenz, A. S. IQ and ability in art. *Sch. & Soc.*, 1928, **27**, 489–490.

166. Lewin, K. *A dynamic theory of personality*. New York and London: McGraw-Hill, 1935.

167. Lewis, N. D. C. Graphic art productions in schizophrenia (dementia praecox). *Proc. Assn. Res. nerv. ment. Dis. for 1925*, **5**, 344–368. New York: Paul B. Hoeber, Inc., 1928.

168. Liese, E. Die freihandzeichnen bei kindlicher Selbständigkeit und unter Berücksichtgung des kindlichen Vermögens. (Freehand drawing and self-reliance in children with special reference to natural ability.) Langensalza: Beltz, 1931. (*Psychol. Abstr.*, 1933, **7**, No. 2084.)

169. Liss, E. Play techniques in child analysis. *Amer. J. Orthopsychiat.*, 1936, **6**, 17–22.

170. Loch, M. Über Eidetik und Kinderzeichnung. (Concerning eidetic imagery and children's drawings.) *Arb. a.d. Psychol., Inst. d. München*, 1931, No. 2. (*Psychol. Abstr.*, 1933, **7**, No. 328.)

171. Lord, E., & Wood, L. Diagnostic values in a visuo-motor test. *Amer. J. Orthopsychiat.*, 1942, **12**, 414–428.

172. Löwenfeld, V. *The nature of creative activity*. New York: Harcourt, Brace and Co., 1939.

173. Löwenfeld, V. Self-adjustment through creative activity. *Amer. J. ment. Def.*, 1941, **45**, 366–373.

174. Löwenfeld, V. *Creative and mental growth*. New York: Macmillan Co., 1947.

175. Luquet, G. H. Le réalisme intellectuel dans l'art primitif. II. Le rendu de visible. (Intellectual realism in primitive art. II. Rendition tion of the visible.) *J. de Psychol.*, 1927, **24**, 888–927.

176. Luquet, G. H. L'évolution du dessin enfantin. (The evolution of children's drawings.) *Bull. Soc. Binet*, 1929, **29**, 145–163.

177. Luquet, G. H. *L'art primitif.* (Primitive art.) Paris: Doin, 1930. (*Psychol. Abstr.*, 1931, **5**, No. 1569.)

178. McAdory, M. The construction and validation of an art test. *Teach. Coll. Contr. Educ.*, No. 383, New York: Teachers' College, Columbia Univ., Bureau of Pub., 1929.

179. McCarthy, Dorothea. A study of the reliability of the Goodenough Test of intelligence. *J. Psychol.*, 1944, **18**, 201–216.

180. McCloy, W. Creative imagination in children and adults. *Psychol. Monogr.*, 1939, **51**, 88–102.

181. McCurdy, H. G. Group and individual variability on the Goodenough Draw-a-Man Test. *J. educ. Psychol.*, 1947, **38**, 428–436.

182. McElwee, Edna W. The reliability of the Goodenough intelligence test used with sub-normal children fourteen years of age. *J. appl. Psychol.*, 1932, **16**, 217–218.

183. McElwee, Edna W. Profile drawings of normal and subnormal children. *J. appl. Psychol.*, 1934, **18**, 599–603.

184. MacGregor, Gordon. *Warriors without weapons*. Chicago: Univ. of Chicago Press, 1946.

185. McHugh, G. Changes in Goodenough IQ at the public school kindergarten level. *J. educ. Psychol.*, 1945, **36**, 17–30.

186. McHugh, G. Relationship between the Goodenough Drawing a Man Test and the 1937 Revision of the Stanford-Binet Test. *J. educ. Psychol.*, 1945, **36**, 119–124.

187. McIntosh, Janette R. An inquiry into the use of children's drawings as a means of psychoanalysis. *Brit. J. educ. Psychol.*, 1939, **9**, 102–103.

188. McIntosh, Janette R., & Pickford, R. W. Some clinical and artistic aspects of a child's drawings. *Brit. J. med. Psychol.*, 1943, **19**, 342–362.

189. Machover, Karen. *Personality projection in the drawing of the human figure.* Springfield, Ill.: C. C Thomas, 1949.

190. Manuel, H. T., & Hughes, Lois S. The intelligence and drawing ability of young Mexican children. *J. appl. Psychol.*, 1932, **16**, 382–387.

191. Marcucci, H. Fanciulli che disegnano. (Children who design.) *Rass. dell'istruz. artistica*, 1932, **3**, 171–180. (*Psychol. Abstr.*, 1933, **7**, No. 3019.)

192. Martin, H. Stile und Stilwandlungsgesetze der Kinderzeichnung, nachgewiesen an den Menschenzeichnungen der Volksschulkinder. (Style and laws of change in style in children's drawing of a man by folks school children.) *Vjsch. Jugendk.*, 1932, **2**, 211–226. (*Psychol. Abstr.*, 1933, **7**, No. 2086.)

193. Martin, H. Die Motivwahl und ihr Wandel in der freien Zeichnung des Grundschulkindes. (Choice of subject and its changes in the free drawings of elementary school children.) *Z. pädag. Psychol.*, 1939, **40**, 231–241.

194. Meier, N. C. Special artistic talent. *Psychol. Bull.*, 1928, **25**, 265–271.

195. Meier, N. C. (Ed.) Studies in the psychology of art. *Psychol. Monogr.*, 1933, **45**, No. 1.

196. Meier, N. C. Studies in the foundations of artistic aptitude. *Proc. Iowa Acad. Sci.*, 1933, **40**, 195–196.

197. Meier, N. C. Art ability without instruction or environmental background: case study of Loran Lockhart. *Psychol. Monogr.*, 1936, **48**, No. 1, 155–163.

198. Meier, N. C. Factors in artistic aptitude: final summary of a ten-year study of a special ability. *Psychol. Monogr.*, 1939, **51**, No. 5, 140–158.

199. Meier, N. C. Recent research in the psychology of art. *Yearb. nat. Soc. Stud. Educ.*, 1941, **40**, 379–400.

200. Meier, N. C., & Seashore, C. E. *The Meier-Seashore Art Judgment Test*, 1929. Revised, 1940 by N. C. Meier under title *The Meier Art Test*. Iowa City, Iowa: Bureau of Educ. Res. and Serv., 1940.

201. Meloun, J. Does drawing skill show in handwriting? *Character & Pers.*, 1935, **3**, 194–213.

202. Menzel, E. W. The Goodenough Intelligence Test in India. *J. appl. Psychol.*, 1935, **19**, 615–624.

203. Merry, Ruth C. Art talent and racial background. *J. educ. Res.*, 1938, **32**, 17–22.

204. Merzbach, A. Symbolische Selbstzeichnungen aus der Psychose eines Jugendlichen und ihre Verwertbarkeit. (Symbolical self-pictures during a psychotic attack in an adolescent and their interpretation.) *Z. ges. Neurol. Psychiat.*, 1930, **127**, 240–251.

205. Miller, J. Intelligence testing by drawings. *J. educ. Psychol.*, 1938, **29**, 390–394.

206. Mira, E. Myokinetic psychodiagnosis: a new technique for exploring the conative trends of personality. *Proc. roy. Soc. Med.*, 1940, **33**, 173–194.

207. Mitchell, W. W. *Drawing aptitude test.* Bloomington, Ill.: McKnight and McKnight, 1940. (Not seen.)

208. Mohr, P. Das künstlerische Schaffen Geisteskranker und seine Beziehungen zum Verlauf der Krank-

heit. (The artistic work of psychotics and its relation to the course of the disease.) *Schweiz. Arch. Neurol. Psychiat.*, 1940, **45**, 427–446. (*Psychol. Abstr.*, 1941, **15**, No. 2625.)

209. Morgenstern, L. L'art plastique et le problème de l'espace. (Plastic art and the problem of space.) *J. de Psychol.*, 1931, **28**, 473–480.

210. Morgenstern, S. Le symbolisme et la valeur psychoanalytique des dessins infantiles. (Symbolism and the psychoanalytic value of children's drawings.) *Rev. franç. Psychanal.*, 1939, **11**, 39–48. (*Psychol. Abstr.*, 1939, **13**, No. 6548).

211. Mott, Sina M. The development of concepts. *J. genet. Psychol.*, 1936, **48**, 199–214.

212. Mott, Sina M. The development of concepts: a study of children's drawings. *Child Develpm.*, 1936, **7**, 144–148.

213. Mott, Sina M. The growth of an abstract concept. *Child Develpm.*, 1939, **10**, 21–25.

214. Mott, Sina M. Muscular activity as an aid in concept formation. *Child Develpm.*, 1945, **16**, 97–109.

215. Munro, T., Lark-Horovitz, Betty, & Barnhart, E. N. Children's art abilities: Studies in the Cleveland Museum of Art. *J. exp. Educ.*, 1942, **11**, 97–155.

216. Munroe, Ruth, Lewinson, Thea S., & Waehner, Trude S. A comparison of three projective techniques. *Character & Pers.*, 1944, **13**, 1–21.

217. Naber, Louise. Relation of child art to primitive art. *Proc. nat. Educ. Assn.*, 1937, **75**, 208–209.

218. Nakae, J. [Children's ways of representing a solid body.] *Kyôiku Shinri Kenkyû*, 1939, **14**, 105–117. (*Psychol. Abstr.*, 1940, **14**, No. 3280.)

219. Napoli, P. J. Finger painting and personality diagnosis. *Genet. Psychol. Monogr.*, 1946, **34**, 129–231.

220. Napoli, P. J. Finger painting: a bibliography. *Sch. Arts*, 1947, **46**, 208–210.

221. Napoli, P. J., & Harris, W. W. Finger painting for the blind. *J. Psychol.*, 1948, **25**, 185–196.

222. National Society for the Study of Education. *Art in American Life and Education.* Fortieth Yearbook, Bloomington, Ill.: Public School Pub. Co., 1941.

223. Naumburg, Margaret. The drawings of an adolescent girl suffering from anxiety hysteria with amnesia. *Psychiat. Quart.*, 1941, **18**, 197–224.

224. Naumburg, Margaret. Children's art expression and war. *Nerv. Child*, 1942, **2**, 360–373.

225. Naumburg, Margaret. A study of the art expression of a behavior problem boy as an aid to diagnosis and therapy. *Nerv. Child*, 1944, **3**, 277–319.

226. Naumburg, Margaret. The psychodynamics of the art expression of a boy patient with tic syndrome. *Nerv. Child*, 1945, **4**, 374–409.

227. Naumburg, Margaret. Fantasy and reality in the art expression of behavior problem children. In Lewis, N. D. C., and Pacella, B. L., *Modern trends in child psychiatry.* New York: International Univ. Press, 1945. Pp. 193–218.

228. Naumburg, Margaret. A study of the psychodynamics of the art work of a nine-year-old behavior problem boy. *J. nerv. ment. Dis.*, 1945, **101**, 28–64.

229. Naumburg, Margaret. A study of the art work of a behavior problem boy as it relates to ego development and sexual enlightenment. *Psychiat. Quart.*, 1946, **20**, 74–112.

230. Naumburg, Margaret. Studies of the "free" art expression of be-

havior problem children and adolescents as a means of diagnosis and therapy. *Nerv. and ment. Dis. Monogr.*, No. 71. New York: Coolidge Foundation, 1947.

231. Neubauer, V. Zur Entwicklung der dekorativen Zeichnung. (The development of decorative drawing.) *Z. angew. Psychol.*, 1931, **39**, 273–325. (*Psychol. Abstr.*, 1931, **5**, No. 4266.)

232. Oakley, C. A. The interpretation of children's drawings. *Brit. J. Psychol.*, 1931, **21**, 256–270.

233. Oakley, C. A. Drawings of a man by adolescents. *Brit. J. Psychol.*, 1940, **31**, 37–60.

234. Oberlin, D. S. Children who draw. *Delaware St. med. J.*, 1938, **10**. (*Psychol. Abstr.*, 1938, **12**, No. 5609.)

235. Okerbloom, Elizabeth C. Hoyt Sherman's experimental work in the field of visual form. *College Art J.*, 1944, **3**, 143–147.

236. Østlyngen, E. [On the direction of drawings.] *Studia Psychol. Paedagog., Lund.*, 1948, **2**, 206–210. (*Psychol. Abstr.*, 1948, **22**, No. 4322.)

237. Paget, G. W. Some drawings of men and women made by children of certain non-European races. *J. roy. Anthrop. Inst.*, 1932, **62**, 127–144.

238. Parlog, C. *Psihologia desemnului.* (*Psychology of drawing.*) Cluj, Roumania: Inst. de Psihol. Univ. Cluj, 1932. (*Psychol. Abstr.*, 1934, **8**, No. 3195.)

239. Paulsson, G. The creative element in art. *Scand. Sci. Rev.*, 1923, **2**, 11–173.

240. Pearson, G. H. J. Inverted position in children's drawings: report of two cases. *J. nerv. ment. Dis.*, 1928, **68**, 449–455.

241. Peck, Leigh. An experiment with drawing in relation to the prediction of school success. *J. appl. Psychol.*, 1936, **20**, 16–43.

242. Pelikan, A. G. *The art of the child.* New York: Bruce Publ. Co., 1931. (*Psychol. Abstr.*, 1932, **6**, No. 467.)

243. Peterson, E. G., & Williams, J. M. Intelligence of deaf children as measured by drawings. *Amer. Ann. Deaf*, 1930, **75**, 273–290.

244. Pfister, H. O. Farbe und Bewegung in der Zeichnung Geisteskranker. (Color and movement in the drawings of mental patients.) *Schweiz. Arch. Neurol. Psychiat.*, 1934, **34**, 325–365.

245. Phillips, E., & Stromberg, E. A comparative study of finger-painting performance in detention home and high school pupils. *J. Psychol.*, 1948, **26**, 507–515.

246. Pintner, R. Artistic appreciation among deaf children. *Amer. Ann. Deaf*, 1941, **86**, 218–224.

247. Piotrowska, I., & Sobeski, M. The primitive. *J. Aesthet.*, 1941, **1**, 12–21.

248. Portocarrero de Linares, Graciela-Vera. El desarrollo del dibujo imatativo en la población femenina de Lima. (The modification of imitative drawing in the feminine population of Lima.) *Bol. Inst. psicopedag. nac. Lima*, 1948, **7**, 47–148.

249. Powel, Lydia. *The art museum comes to the school.* New York: Harper and Bros., 1944.

250. Prudhomme, M. *Le dessin de l'enfant.* (The child's drawing.) Paris: Presses Universitaires de France, 1947.

251. Rabello, S. Caractericas do desenho infantil. Contribução para o estudo psychologico da criança brasileira. (Characteristics of children's drawings. A contribution to the psychological study of Brazilian children.) *Bol. Direct. tech. Educ., Pernambuco*, 1932, **2**,

15–78. (*Psychol. Abstr.*, 1935, **9**, No. 3035.)

252. RAY, ANDRÉ. Épreuves de dessin; témoins du développement mental. (Drawing tests as evidence of mental development.) *Arch. Psychol., Genève*, 1946, **31**, 369–380. (*Psychol. Abstr.*, 1947, **21**, No. 4178.)

253. RAY, ANDRÉ. Épreuves de dessin; témoins du développement mental: II. Reproduction d'objets simples selon différents points de vue. (Drawing tests as evidence of mental development: II. Reproduction of simple objects from different points of view.) *Arch. Psychol., Genève.*, 1947, **32**, 145–160. (*Psychol. Abstr.*, 1948, **22**, No. 2828.)

254. READ, H. E. *Education through art.* London: Faber and Faber, 1943.

255. REITMANN, F. Facial expression in schizophrenic drawings. *J. ment. Sci.*, 1939, **85**, 264–272.

256. RICE, CHARLOTTE. Excellence of production and types of movements in drawing. *Child Develpm.*, 1930, **1**, 1–14.

257. RODGERS, FRANCES. Variation in the aesthetic environment of artistic and non-artistic children. *Psychol. Monogr.*, 1933, **45**, No. 1, 95–107.

258. ROSTOHAR, M. Studiez vyvojové psychologie. Dil. I. (Studies in developmental psychology.) Brno: Opera Facult. Philosoph. Universit. Masarykian. Brunensis, Cisto 25, 1928. (*Psychol. Abstr.*, 1930, **4**, No. 805.)

259. ROUMA, G. Dessins d' indiens quitchouas et aymaras. La graphisme et l'expression graphique. (Drawings of the Quechua and the Aymara Indians. Writing and graphic expression.) *Sem. univ. Pédag., Univ. libre Brux.*, 1935, **1**, 133–147. (*Psychol. Abstr.*, 1936, **10**, No. 4173.)

260. RUESCH, J., & FINESINGER, J. E. The relation of the Rorschach color response to the use of color in drawings. *Psychosom. Med.*, 1941, **3**, 370–388.

261. RUSSELL, R. W. The spontaneous and instructed drawings of Zuni children. *J. comp. Psychol.*, 1943, **35**, 11–15.

262. SAUNDERS, A. W. The stability of artistic aptitude at the childhood level. *Psychol. Monogr.*, 1936, **48**, 126–154.

263. SCHAEFER-SIMMERN, H. *The unfolding of artistic activity, its basis, processes, and implications.* Berkeley: Univ. California Press, 1948.

264. SCHEER, G. Ausdrucksfaktoren in der zeichnerischen Entwicklung des Pubertätsalters. (Expressive factors in the development of drawing during adolescence.) *Arch. ges. Psychol.*, 1934, **92**, 1–25.

265. SCHILDER, P. The child and the symbol. *Scientia*, 1938, **64**, 21–26.

266. SCHILDER, P., & LEVINE, E. L. Abstract art as an expression of human problems. *J. nerv. ment. Dis.*, 1942, **95**, 1–10.

267. SCHLIEBE, G. Erlebnismotorik und zeichnerischer (physiognomischer) Ausdruck bei Kindern und Jugundlichen. (Zur Psychogenese der Ausdruckesgestaltung) (Kinesthetic experience and graphic [physiognomical] expression among children and adolescents. [On the psychogenesis of the expression of form.]) *Z. Kinderforsch.*, 1934, **43**, 49–76.

268. SCHMIDL-WAEHNER, TRUDE. Formal criteria for the analysis of children's drawings. *Amer. J. Orthopsychiat.*, 1942, **12**, 95–103.

269. SCHUBE, P., & COWELL, J. Art of psychotic persons. *Arch. Neurol. Psychiat., Chicago*, 1939, **41**, 709–720.

270. SCHUBERT, ANNA. Drawings of Orotchen children and young peo-

ple. *J. genet. Psychol.*, 1930, **37**, 231–244.

271. SEEMAN, E. Development of the pictorial aptitude in children. *Character & Pers.*, 1934, **2**, 209–221.

272. SEEMAN, E., & SAUDEK, R. The self-expression of identical twins in handwriting and drawing. *Character & Pers.*, 1932, **1**, 91–128.

273. SHAFFER, L. F. Children's interpretations of cartoons. *Teach. Coll. Contr. Educ.*, No. 429. New York: Bureau of Pub., Teachers' College, Columbia Univ., 1930.

274. SHAW, RUTH F. *Finger painting.* Boston: Little, Brown and Co., 1934. (*Psychol. Abstr.*, 1935, **9**, No. 5969.)

275. SHIRLEY, MARY, & GOODENOUGH, FLORENCE L. A survey of intelligence of deaf children in Minnesota schools. *Amer. Ann. Deaf*, 1932, **77**, 238–247.

276. SIGG-BOEDDINGHAUS, M. Zeichnen und Malen und ihre Bedeutung für die seelische Entwicklung des Kindes. (Drawing and painting; their significance for the mental development of the child.) *Psychol. Rundschau*, 1929, **1**, 222–229; 324–330. (*Psychol. Abstr.*, 1930, **4**, Nos. 2185, 2470.)

277. SLOCHOWER, MURIEL Z. Experiments on dimensional and figural processes in the clay and pencil reproductions of line figures by young children: I. Dimension. *J. genet. Psychol.*, 1946, **69**, 57–75.

278. SLOCHOWER, MURIEL Z. Experiments on dimensional and figural problems in the clay and pencil reproductions of line figures by young children: II. Shape. *J. genet. Psychol.*, 1946, **69**, 77–95.

279. SORGE, S. Neue Versuche über die Wiedergabe abstraktes optischer Gebilde. (A new experiment on the reproduction of abstract visual forms.) *Arch. ges. Psychol.*, 1940, **106**, 1–89.

280. SPIELREIN, S. Kinderzeichnungen bei offnen und geschlossenen Augen; Untersuchungen über die unterschwelligen kinästhetischen Vorstellungen. (Children's drawings with the eyes open and shut: researches on kinesthetic images below the threshold of consciousness.) *Imago*, 1931, **17**, 359–391. (*Psychol. Abstr.*, 1932, **6**, No. 2966.)

281. SPOERL, DOROTHY T. Personality and drawing in retarded children. *Character & Pers.*, 1940, **8**, 227–239.

282. SPOERL, DOROTHY T. The drawing ability of mentally retarded children. *J. genet. Psychol.*, 1940, **57**, 259–277.

283. SPOERL, DOROTHY T. A note on the Anastasi-Foley cultural interpretation of children's drawings. *J. soc. Psychol.*, 1941, **13**, 187–192.

284. SPRINGER, N. N. A comparative study of the intelligence of a group of deaf and hearing children. *Amer. Ann. Deaf*, 1938, **83**, 138–152.

285. SPRINGER, N. N. A study of the drawings of maladjusted and adjusted children. *J. genet. Psychol.*, 1941, **58**, 131–138.

286. STEGGERDA, M. The McAdory Art Test applied to Navaho Indian children. *J. comp. Psychol.*, 1936, **22**, 283–285.

287. STEGGERDA, M. *Maya Indians of Yucatan.* Washington, D.C.: Carnegie Institution of Washington, 1941.

288. STOLZ, ALBERTA, & MANUEL, H. T. The art ability of Mexican children. *Sch. & Soc.*, 1931, **34**, 379–380.

289. TAYLOR, W. S. A note on cultural determination of free drawings. *Character & Pers.*, 1944, **13**, 30–36.

290. TELFORD, C. W. Test performance of

full and mixed-blood North Dakota Indians. *J. comp. Psychol.*, 1932, **14**, 123–145.

291. Thiel, G. Eine Untersuchung von Kinderzeichnungen taubstummer Schüler. (A research on the drawings of deaf children.) *Z. Kinderforsch.*, 1927, **33**, 138–176.

292. Thompson, Laura, & Joseph, Alice. *The Hopi way.* Chicago: Univ. of Chicago Press, 1945.

293. Tiebout, Carolyn. The psychophysical functions differentiating artistically superior from artistically inferior children. *Psychol. Monogr.*, 1933, **45**, No. 1, 108–133.

294. Tiebout, Carolyn. The measurement of psychophysical functions characteristic of extreme degrees of artistic capacity. *Proc. Iowa Acad. Sci.*, 1933, **40**, 197.

295. Tiebout, Carolyn. The measurement of quality in children's painting by the scale method. *Psychol. Monogr.*, 1936, **48**, No. 1, 85–94.

296. Tiebout, Carolyn, & Meier, N. C. Artistic ability and general intelligence. *Psychol. Monogr.*, 1936, **48**, No. 1, 95–125.

297. Todd, Jessie M. *Drawing in the elementary school.* Chicago: Univ. Chicago Press, 1931.

298. Tomlinson, R. R. *Picture making by children.* London: The Studio, Ltd., 1934.

299. Tomlinson, R. R. *Children as artists.* London: The King Penguin Books, 1944.

300. Traube, T. Le valeur diagnostique des dessins des enfants difficiles. (The diagnostic value of drawings of problem children.) *Arch. Psychol., Genéve*, 1937, **26**, 285–309.

301. Vaughn, C. L., & Hoose, Elizabeth S. Special abilities in a mentally deficient boy. *J. Psycho-Asth.*, 1936, **41**, 197–207.

302. Vedder, R. Over het copieren von eenvoudige geometrische figuren door oligophrenen en jonge kinderen. (The copying of simple geometric figures by feebleminded and young children.) Amsterdam: N. V. Noord-Hollandsche Uitgevers Maatschappi, 1939. (*Psychol. Abstr.*, 1940, **14**, No. 4633.)

303. Viola, W. *Child art.* London: Univ. of London Press, 1942.

304. Vogelsang, H. Aus der Ideenwelt jugendlicher Rechtsbrecher. (The mental world of the young lawbreaker.) *Z. Jugendk.*, 1934, **4**, 120–125.

305. Voss, Mildred D. A study of conditions affecting the functioning of the art appreciation process at the child-level. *Psychol. Monogr.*, 1936, **48**, No. 1, 1–39.

306. Voss, Mildred D. The validity and reliability of a modified form of the McAdory Art Test for use at lower grade levels. *Psychol. Monogr.*, 1936, **48**, No. 1, 68–84.

307. Vydra, J. Stammesunterschiede in der Begabung des Kindes für bildende Kunst. (Racial differences in children's talent for pictorial art.) 4 *Versamml. f. Kinderforsch.*, Bratislava, 1932, 311–314. (*Psychol. Abstr.*, 1933, **7**, No. 4091.)

308. Waehner, Trude S. Interpretations of spontaneous drawings and paintings. *Genet. Psychol. Monogr.*, 1946, **33**, 3–70.

309. Wartegg, E. Gestaltung und Charakter. (Organization and character.) *Beih. Z. angew. Psychol.*, 1939, No. 84.

310. Wartegg, E. *Gestaltung und Charakter. Ausdrucksdeutung zeichnerischer Gestaltung und Entwurf einer charakterologischen Typologie.* (Form and character. Interpretation of expression from drawings and an outline of characterological typology.) Leipzig: J. A. Barth,

1939. (*Psychol. Abstr.*, 1941, **15**, No. 4720.)

311. WECHSLER, D., & HARTOGS, R. The clinical measurement of anxiety. *Psychiat. Quart.*, 1945, **19**, 618–635.
312. WHITE, ROBERTA, & JOHNSON, BUFORD. Children's choices in modern art. *Child Develpm.*, 1930, **1**, 347–349.
313. WIGGIN, R. *Developmental stages of drawing expression.* (Mimeographed) Author: College Park, Md., 1948.
314. WILLIAMS, J. H. Validity and reliability of the Goodenough Intelligence Test. *Sch. & Soc.*, 1935, **41**, 653–656.
315. WILLIAMS, J. N. Interpretation of the drawings made by maladjusted children. *Virginia med. Mon.*, 1940, **67**, 533–538.
316. WILLIAMS, M. L. The growth of intelligence as measured by the Goodenough Drawing Test. *J. appl. Psychol.*, 1930, **14**, 239–256.
317. WINTSCH, J. Le dessin comme témoin du développement mental. (Drawing as evidence of mental development.) *Z. Kinderpsychiat.*, 1935, **2**, 33–44; 69–83.
318. WITMER, HELEN L. (Ed.) *Psychiatric interviews with children.* New York: Commonwealth Fund, 1946.
319. WOLFF, W. "Graphometry": a new diagnostic method. *Psychol. Bull.*, 1942, **39**, 465.
320. WOLFF, W. Projective methods for personality analysis of expressive behavior in preschool children. *Character & Pers.*, 1942, **10**, 309–330.
321. WOLFF, W. *The personality of the preschool child; the child's search for his self.* New York: Grune and Stratton, 1946.
322. WOLFF, W. *Diagrams of the unconscious.* New York: Grune and Stratton, 1948.
323. WOOD, L., & SHULMAN, E. The Ellis visual designs test. *J. educ. Psychol.*, 1940, **31**, 591–602.
324. YEPSON, L. N. The reliability of the Goodenough Drawing Test with feebleminded subjects. *J. educ. Psychol.*, 1929, **20**, 448–451.
325. ZACHRY, CAROLINE B. The role of art in personality development. *Ment. Hyg.*, (N.Y.), 1933, **17**, 51–58.
326. ZAMBALDI, I. La scuola attiva e il metodo d'insegnamento. Il disegno. (The active school and the method of teaching. Design.) *Riv. pedag.*, 1934, **27**. (*Psychol. Abstr.*, 1935, **9**, No. 2005.)
327. ZESBAUGH, HELEN A. *Children's drawings of the human figure.* Chicago: Univ. of Chicago Press, 1934.
328. ZILAHI-BEKE, A. Zusammenhänge zwischen Kunst- und Charakterentwicklung. (Relationship between art and character development.) *Int. Z. Indiv.-Psychol.*, 1931, **9**, 51–60.
329. ZILGITT, HELEN H. Relationships in child development. *Design.* 1935, **37**, 26–27.
330. ZIMMERMAN, J., & GARFINKLE, L. Preliminary study of the art productions of the adult psychotic. *Psychiat. Quart.*, 1942, **16**, 313–318.

Received March 21, 1950.

ON THE VALIDITY OF INFERENCES DRAWN FROM THE FITTING OF POISSON AND NEGATIVE BINOMIAL DISTRIBUTIONS TO OBSERVED ACCIDENT DATA

J. S. MARITZ

National Institute for Personnel Research, South African Council for Scientific and Industrial Research

I. Introduction

In a recent article Mintz and Blum (3) pointed out that the so-called "Method of Percentages" sometimes used to expose accident proneness is fallacious. With this statement the writer is in agreement, but it is felt at the same time that the method proposed by the authors of the above-mentioned article may lead to erroneous conclusions if the underlying assumptions are not met with in practice. Mintz and Blum accept the satisfactory representation of a set of accident data by the Negative Binomial distribution as sufficient evidence that there exists accident proneness in the group from which the data originated. Conversely, a good Poisson "fit" should indicate absence of proneness. That such an interpretation of the Negative Binomial law may not always be in accordance with the facts has recently been stated by Greenwood (2), one of the pioneers in statistical analysis of accident data.

It is known that a composite of only two Poisson distributions can well be represented by the Negative Binomial law so that in this article attention will be given to Poisson and Negative Binomial distributions only (5). It is the object of this paper to show, firstly, that a perfect Poisson "fit" does not exclude proneness, and secondly, that a good Negative Binomial "fit" need not indicate accident proneness.

The most direct method of establishing proneness in a group of people, all of whom ought to be exposed to the same environmental risk, consists of splitting a lengthy period of observation into two and correlating the frequency of accidents per individual for these two periods. This statistical technique is nearest to the psychological definition of accident proneness. As accident distributions are far from normal, the resulting product-moment coefficient r can only be interpreted as a sloping constant of the regression lines, these being the lines of best fit to the data in the sense of least squares. Ordinarily, however, proneness is established from univariate distributions by comparing the variance of the accident data with the mean. The justification of this procedure may be found in the mathematical relationship of the correlation coefficient between the number of accidents in two sucessive periods and the parameters of the Negative Binomial law (Newbold, 4).

II. Poisson Distribution

A well known discrete bivariate law with marginal distributions of the Poisson type and non-zero correlation is given by Aitken (1) as:

$$p(x, y) = e^{-\lambda_1-\lambda_2+\lambda_3}\frac{(\lambda_1-\lambda_3)^x(\lambda_2-\lambda_3)^y}{x!y!}\left\{1+\frac{xy}{(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)}\frac{\lambda_3}{1!} + \frac{x(x-1)y(y-1)}{(\lambda_1-\lambda_3)^2(\lambda_2-\lambda_3)^2}\frac{\lambda_3^2}{2!}+\cdots\right\}. \qquad [1]$$

For the marginal distributions we find by summation from equation [1],

$$\left.\begin{aligned} p(x) &= e^{-\lambda_1}\frac{\lambda_1^x}{x!} \\ p(y) &= e^{-\lambda_2}\frac{\lambda_2^y}{y!} \end{aligned}\right\}. \qquad [2]$$

The following example has been constructed to illustrate some properties of the above bivariate Poisson distribution. The frequencies of Table I were calculated, using the following values for the parameters:

$$\lambda_1 = \lambda_2 = \cdot 5,$$
$$\lambda_3 = \cdot 4.$$

TABLE I

An Example Illustrating Properties of the Bivariate Poisson Distribution

Period 1	Period 2					
	0	1	2	3	4	
4			.1	.3	.3	.7
3	.1	.6	2.3	3.1	.3	6.4
2	1.4	11.1	23.1	2.3	.1	38.0
1	27.4	112.5	11.1	.6		151.6
0	274.4	27.4	1.4	.1		303.3
	303.3	151.6	38.0	6.4	.7	500.0

The marginal distributions are, therefore,

$$p(x) = p(y),$$

$$p(x) = e^{-.5}\frac{(\cdot 5)^x}{x!}\cdot$$

Calculating the correlation coefficient from the data of Table I leads to $r = .80$. It may be shown from equation [1] that the correlation coefficient of the bivariate population is

$$\rho = \frac{\lambda_3}{\sqrt{\lambda_1\lambda_2}} = \frac{.4}{\sqrt{.5 \times .5}} = .8. \qquad 3]$$

Furthermore it may be shown from equation [1] that the regressions of x on y and of y on x are rectilinear, hence the interpretation of the correlation coefficient as a sloping constant of the regression lines. The regression lines pertaining to the data in Table I are graphed in Fig. 1.

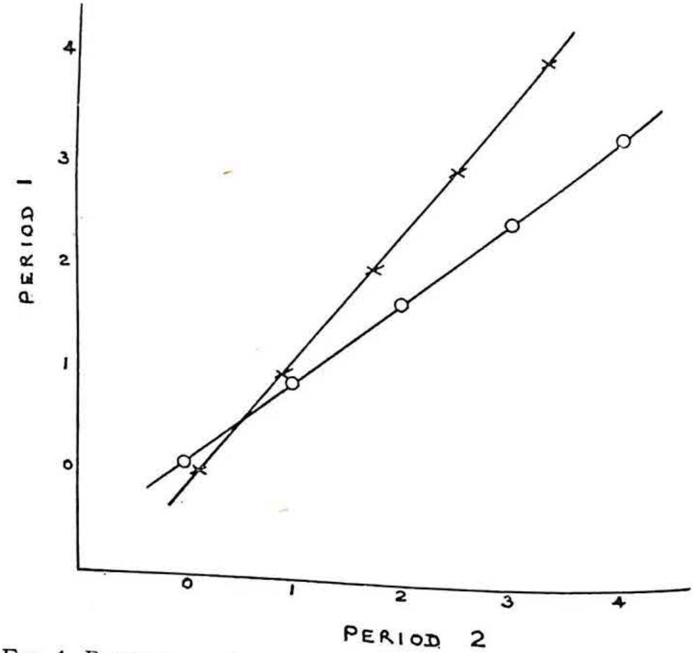


FIG. 1. REGRESSION LINES BASED UPON THE DATA IN TABLE I

The above example shows how wrong our conclusions would have been had we analyzed the distribution of accidents within a single period by the usual procedure of comparing the mean with the variance.

It might be interesting to see what happens if we combine the two successive periods into one longer single period. By adding the frequencies in Table I along the lines 0-0, 1-1, 2-2, etc., we arrive at the following univariate distribution.

No. of accidents	0	1	2	3	4	5	6	7	8	Total
Frequencies	274.4	54.8	115.3	22.4	24.3	4.6	3.3	.6	.3	500.0

mean = 1.0
variance = 1.8

This distribution is not Poisson any more, since variance > mean.

The theoretical distribution, the derivation of which will be published elsewhere, is given by

$$p(z) = e^{-\lambda_1-\lambda_2+\lambda_3}(\lambda_1 + \lambda_2 - 2\lambda_3)^z \sum_{r=0}^{\leqq z/2} \frac{\lambda_3{}^r}{(\lambda_1 + \lambda_2 - 2\lambda_3)^{2r}} \frac{1}{r!(z - 2r)!}, \qquad [4]$$

$$\left.\begin{aligned} &\text{with mean} && = \lambda_1 + \lambda_2 = 1.0 \\ &\text{and variance} && = \lambda_1 + \lambda_2 + 2\lambda_3 = 1.8 \end{aligned}\right\}. \qquad [5]$$

It appears, therefore, that by taking a long enough period of observation we should eventually find such differences between mean and variance as should be indicative of accident proneness, although it may not be sufficient evidence. However, the correlational approach is to be preferred since in practice we can say very rarely a priori how long such a period should be.

Practical Application

The data in Table II were supplied by Adelstein[1], referring to serious

TABLE II

DATA FROM ADELSTEIN'S STUDY (UNPUBLISHED) OF ACCIDENT RATES OF RAILWAY OPERATIVES

NO. OF ACCIDENTS	OBSERVED FREQUENCIES		THEORETICAL FREQUENCIES		χ^2
0	44		39.7		.466
1	42		44.6		.152
2	21		25.0		.640
3	9	15	9.4	12.7	.416
4	4		2.6		
5	2		.6		
6	0		.1		
TOTAL NO.	122		122.0		1.674

$d.f. = 2$
$P = .44$

---

[1] Unpublished study.

accidents incurred during a period of five years by 122 railway operatives. The Poisson distribution was fitted to these data and we cannot reject the hypothesis that this sample was drawn from a Poisson population of accidents. We should, therefore, conclude that the accidents were following a pure chance pattern especially since the length of the observational period was considerable. The observed and fitted Poisson distributions are given in Table II.

The accident distribution of the same 122 people for a subsequent period of six years was also available. The chi-square test again revealed that we could not reject the hypothesis of a pure chance distribution. The bivariate distribution of these accidents is given in Table III.

TABLE III

BIVARIATE ACCIDENT DISTRIBUTION BASED UPON ADELSTEIN'S STUDY

PERIOD 1 (YEARS 1–5)	PERIOD 2 (YEARS 6–11)								TOTAL No.
NO. OF ACCIDENTS	0	1	2	3	4	5	6	7	
0	21	14	8	1					44
1	17	12	8	3	1			1	42
2	6	9	2	2	2				21
3	1	1	3	3	1				9
4	1	3							4
5				2					2
6									
7									
TOTAL No.	46	39	21	11	4			1	122

The product-moment coefficient as computed from the above table is $r = .29$, which refutes the previous conclusion of no proneness within the group as derived from the chi-square tests of the observed univariate accident distributions. Equation [1] was fitted to the observed data of Table III and the resulting test of goodness of fit gave, for 7 *d.f.*, $P = .49$. Hence it follows that the above data follow a *correlated* bivariate Poisson distribution.

III. NEGATIVE BINOMIAL DISTRIBUTION

If the hypothesis of a Poisson accident law is refuted by the chi-square test, it is customary to proceed to graduate the data by the Negative Binomial distribution. In this section it will be shown that a good Negative Binomial "fit" need not indicate accident proneness within the group.

Newbold (4) gives the relation between the correlation of the accidents within two equal successive time periods and the parameters of the univariate Negative Binomial distribution (representing the frequency of accidents in the combined time period) as

$$\rho = \frac{t}{2p + t}, \tag{6}$$

where the Negative Binomial law of the total period is defined as

$$p(z) = \left(\frac{p}{p + t}\right)^p \frac{\Gamma(p + z)}{\Gamma(p)z!} \left(\frac{t}{p + t}\right)^z. \tag{7}$$

The above two equations give us the mathematical rationale for using a good Negative Binomial "fit" as an indication of accident proneness, since a good Negative Binomial "fit" by itself would not suffice to indicate proneness, which by definition is associated closely with correlation. It must be emphasized that the above reasoning is only valid if the parameter p is identical for the combined period and for the two individual shorter periods which are correlated. One may, however, construct a bivariate Negative Binomial distribution of zero correlation; i.e.

$$\left.\begin{aligned} f(x, y) &= p(x)p(y) \\ \text{where} \qquad p(x) &= \left(\frac{p}{p + t}\right)^{p/2} \frac{\Gamma(p/2 + x)}{\Gamma(p/2)x!} \left(\frac{t}{p + t}\right)^x \\ p(y) &= \left(\frac{p}{p + t}\right)^{p/2} \frac{\Gamma(p/2 + y)}{\Gamma(p/2)y!} \left(\frac{t}{p + t}\right)^y \end{aligned}\right\}. \tag{8}$$

From equation [8] the distribution law of the accidents incurred in the combined period (variable $z = x + y$) may be deduced; it is identical with equation [7]. From equations [7] and [8] we see that the condition of equal p-constants for the individual as well as for the combined periods is not fulfilled. Hence equation [6] will not hold, and any inference based on Newbold's argument will be erroneous.

A hypothetical example may make the latter point clearer. Suppose that the accidents of one thousand workers were distributed as indicated in the first two columns of Table IV. A Negative Binomial fit of $P = .826$ for 14 degrees of freedom indicates a very good representation of the data.

From the distribution in Table IV we find

$$m_1' = 4.0100,$$
$$s^2 = 19.8959.$$

TABLE IV

HYPOTHETICAL EXAMPLE OF OBSERVED AND EXPECTED ACCIDENT FREQUENCIES

No. of Accidents	f_O	f_E	$f_O - f_E$	$\frac{(f_O - f_E)^2}{f_E}$
0	202	197.6	+4.4	.098
1	161	159.7	+1.3	.010
2	128	128.3	− .3	.001
3	102	102.9	− .9	.008
4	81	82.4	−1.4	.024
5	67	65.9	+1.1	.018
6	52	52.8	− .8	.012
7	41	42.2	−1.2	.034
8	34	33.7	+ .3	.003
9	22	27.0	−5.0	.926
10	17	21.6	−4.6	.980
11	17	17.2	− .2	.002
12	9	13.8	−4.8	1.670
13	14 } 23	11.0 } 19.8	+3.2	.517
14	9	8.8		
15	10 } 20	7.0 } 12.6	+7.4	4.346
16	10	5.6		
17	5 } 13	4.5 } 11.0	+2.0	.364
18	3	3.6		
19	5	2.9		
20	2 } 11	11.5	− .5	.022
21	4			
22	2			
23				
24	2			
25				
26	1			
Total No.	1000	1000.0	0.0	9.035

For the estimates of the parameters p and t in equation [7] we find

$$\hat{t} = m_1' = 4.0100,$$

$$\hat{p} = \frac{m_1'^2}{s^2 - m_1'} = 1.0122.$$

Substitution of these values into equation [6] gives us an estimate for ρ, i.e.

$$r = \frac{4.0100}{2.0244 + 4.0100} = .66.$$

From the satisfactory goodness of fit and the high calculated correlation we should conclude that there exists an appreciable amount of proneness in this particular group.

TABLE V

HYPOTHETICAL BIVARIATE DISTRIBUTION OF 1000 ACCIDENT FREQUENCIES

PERIOD 1 (Y)	PERIOD 2 (X)																									TOTAL No.
No. Acc.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
0	202	80	48	32	22	16	12	9	6	5	4	3	2	2	1	1	1	1		1		1			1	450
1	81	32	19	13	9	6	5	3	3	2	2	1	1		1		1		1							180
2	48	19	11	8	6	4	3	2	2		1		1		1		1									107
3	32	13	8	5	3	3	2	1	1		1	1	1		1											72
4	22	9	5	3	3	2	1	1	1	1		1					1									50
5	17	6	4	2			1		1	2		1														34
6	12	5	3		2			1			1					1										25
7	8	4		2	1		1	1		1																18
8	7	3		1		1					1															13
9	5		2		1		1	1																		10
10	3	2	1	1		1	1																			9
11	3	1		1		1																				6
12	2	1	1	1																						5
13	2		1											1												4
14	1	1	1						1																	4
15	1				1		1																			3
16	1	1	1																							3
17	1			1																						2
18		1			1																					2
19	1																									1
20																										
21	1																									1
22																										
23																										
24	1																									1
TOTAL No.	451	178	105	70	49	34	28	19	15	11	10	7	5	3	4	2	4	1	1	1		1			1	1000

Mintz and Blum (3) proposed to measure the amount of accident proneness by the "percentage of variance attributable to differences in accident liability," the formula being

$$\text{percentage} = \frac{\text{variance} - \text{mean}}{\text{variance}} \times 100.$$

Substituting our values into this expression we have

$$\text{percentage} = 80,$$

from which we again would infer that we deal with a group in which accident proneness is very pronounced.

Supposing now we split the period of observation into two half-periods, we may arrive at a scattergram as shown in Table V (addition along the lines 0-0, 1-1, 2-2, etc. will again lead to the univariate distribution of Table IV).

For the means and standard deviations of the marginal distributions of Table V we find

Period 1	*Period* 2
$\bar{y} = 2.00100$	$\bar{x} = 2.00900$
$s_y^2 = 10.17900$	$s_x^2 = 9.93892$

$$\frac{1}{1000}\sum xy = 3.90900$$

$$r = \frac{-.11101}{10.05824},$$

$$= -.011.$$

The regression lines of the bivariate distribution of Table V are shown in Fig. 2.

From the above correlation we must infer that the accidents were not caused by individual proneness. This result is in striking contrast to the conclusions based on the usual type of analyses.

IV. SUMMARY

It has been shown that the fitting of theoretical Poisson and Negative Binomial distributions to observed accident data alone is not an adequate criterion for the detection of accident proneness.

It is suggested that the direct technique of "correlating consecutive periods" is indispensable.

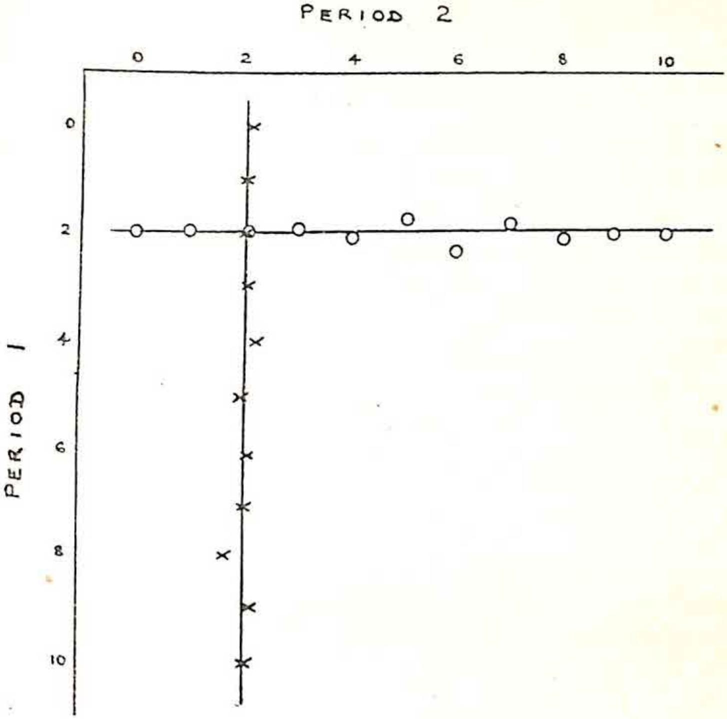


FIG. 2. REGRESSION LINES FOR THE BIVARIATE DISTRIBUTION OF TABLE V

BIBLIOGRAPHY

1. AITKEN, A. C. *Statistical mathematics*, University Mathematical Texts, 2. Edinburgh and London: Oliver and Boyd, 1944.
2. GREENWOOD, M. Discussion of paper: Some statistical aspects of road safety research, by R. J. Smeed. *J. roy. statist. Soc.*, A, 1949, **112**, 25.
3. MINTZ, A., & BLUM, M. L. A re-examination of the accident proneness concept. *J. appl. Psychol.*, 1949, **33**, 195–211.
4. NEWBOLD, E. M. Practical applications of the statistics of repeated events, particularly to industrial accidents. *J. roy. statist. Soc.*, 1927, **90**, 487.
5. PEARSON, K. On certain types of compound frequency distributions in which the components can be individually described by binomial series. *Biometrika*, 1915, **11**, 139–144.

Received February 24, 1950.

BOOK REVIEWS

UNDERWOOD, BENTON J. *Experimental psychology: An introduction.* New York: Appleton-Century-Crofts, 1949. Pp. vii+638. $4.50.

The goal of this volume is a unification of the subject matter and methods of experimental psychology. In the course of fourteen chapters the following topics are considered: discriminal processes, motivation, frustration, conflict, transfer of training, conditioning, multiple response learning, thinking, retention, and, finally, work (which is conceived to involve in a complex manner all of the other processes). The materials are organized from the standpoint of a "dimensional analysis of manipulable variables." Throughout, there is an emphasis on the principles of experimental design, and on some of the more basic principles of scientific logic.

The materials presented are carefully chosen to illustrate both sound and faulty experimental reasoning and execution. For the most part, experimental selections reflect contemporary research interests and include, for example, studies of latent learning, place-learning, and the spread of effect, under theoretical problems of learning; under conflict, Miller's gradient-analysis, Maier's studies of experimental neurosis, and Mowrer's demonstration of "regression" are among the examples included. Experiments with applied significance, when they meet the criteria of good design, are given, such as Berger's studies of perception of form, size, and color combinations of number plates, or Woodward's study of transfer of training in industrial operations. The range of problems considered extends to complex processes of human adjustment investigated in such studies as those of ego-involvement, level of aspiration, or of motivational influences on perception. Each of the major topics is organized around the variables which have been experimentally demonstrated, or which may be supposed to influence the process under consideration. Typically, the materials included in the chapter on thinking are organized around the variables of "massed vs. distributed training," "type of material," "set," and "memory factors." Because the basis of the treatment is methodological, work done under divergent theoretical frameworks, such as Hull's, Tolman's, or Lewin's, is readily assimilated into the text without creating any undue impression of inconsistency.

Four types of experimental design are expounded in detail: Method I which employs random groups, Method II which employs groups matched on performance in advance of introducing the experimental variable, Method III which employs complete counterbalancing of conditions, and Method IV which is based on a systematic randomization of conditions. Variations and applications of each method are discussed, together with criteria for the choice of each of the several methods under given sets of conditions.

The methodological approach is Operationistic, and the treatment gains and loses according to the strengths and weaknesses of this general position. Dimensions of behavior are, wherever possible, to be established by experimental operations. A "dimension" is defined as a stimulus or response scale, some aspect of which is shown to vary reliably in amount. Generalizations tend, under this viewpoint, to be limited to statements of functional relationships between dimensions. However, it is recognized that "a mere correlation of the physical stimulus dimension with the response dimension does not provide a complete picture of the causal factors in behavior." The gap is filled by means of the construct which is defined as "an hypothetical process or state of the organism, the operation of which we infer from observations of the behavior of the organism." Motivation is, according to this view, a state of energization which is inferred from changes in output, or muscular tension exhibited by the organism. Such a construct is viewed as a device for avoiding the necessity for defining particular operations for each "motivational" experiment, making it possible to reduce the many separate operations to a few common factors.

But some of the difficulties involved in carrying through a consistent dimensional analysis become apparent in the course of the writing. In a number of transfer of training phenomena which are contingent on similarity between conditions, the facts simply do not permit a dimensional analysis, as Underwood is careful to state. And in spite of a more or less implicit initial contract that established dimensions will be utilized, it is recognized later that some of the scales must be "improvised." A "motor-verbal continuum" is admittedly a difficult one to dimensionalize because of a lack of comparable units of measurement. The shortcomings of a dimensional analysis become critical, in the reviewer's judgment, in Underwood's attempt to unify the several types of learning—conditioning, multiple-response learning and thinking, which are described as "modal points (presumably on the same continuum) of laboratory operations. . . . We do not assume that there are different *kinds* of learning; the relationships among all tasks used to study learning are too apparent to admit such a distinction *at this time* [italics are Underwood's]. It is a matter for the future to produce the evidence requiring abandonment of this position and establishing the fact that certain distinctions are necessary." The reasoning at this point, the reviewer would contend, is the reverse of accepted scientific logic, which throws the burden of proof for such a priori unification on the theorist. The works of Maier, Jacobsen, Hilgard and Marquis, to cite just a few supports, make it plain that experimentable distinctions do exist between various types of learning situations. Moreover, so far as the reviewer is aware, no generally accepted theoretical resolution of the problem of "levels" of learning has been accomplished. The substantial advances of contemporary behavior theory over the classical Behavioristic position toward learning have resulted from explicit recognition and refinement of distinctions among *types* of stimulating circumstances, *types* of response, and *types* of learned adaptations. The same point should be made with reference to the interesting speculation that motivational variations be conceived as lying on the same continuum with emotional changes.

In general, the descriptions of experiments are accurate and clear. In a few cases the description is possibly too sketchy for complete comprehension, as in the discussions of delayed response or of Maier's reasoning experiments. In the summary of an experiment by Bugelski on derived reinforcement, the sense is lost because it is stated that the experimental and control animals are both trained and tested under the differential conditions. Actually both groups were exposed to the same conditions during acquisition of the habit, but during extinction the derived reinforcing stimulus was omitted for one group and retained for the other. But in view of the large mass of detailed experimental materials summarized, there are very few inaccuracies.

It is the reviewer's opinion that the book offers a set of highly worthwhile background materials for a strong course in experimental psychology. The type of need which it fulfills for the student is seen clearly in the sections dealing with extrapolation from data obtained with animals to human behavior, and in the discussion of guides for scientific generalization, both of which could be reviewed with profit by many of us who enjoy being thought of as experimentalists. Those who have concluded, with the reviewer, that courses in the physical sciences frequently *fail* to imbue students with an understanding of scientific method, and who feel, consequently, that developing skill in the application of scientific modes of thought to complex problems of human nature and social interaction is a task that falls to the psychologist, will be quick to sense the value of Underwood's effort.

Oberlin College. JOHN L. FINAN.

STAGNER, ROSS. *Psychology of personality.* (Rev. Ed.) New York: McGraw-Hill, 1948. Pp. xiii+485. $5.00.

Ross Stagner has completely revised his 1937 text on personality and now presents a longer, more inclusive, more systematic and, in the eyes of the present reviewer, a more gracefully written book.

The new edition, like the old, is organized into four major sections. Section I, *Introduction*, disposes of the definition of personality and devotes two chapters to methods. Section II, *Description of Personality*, includes material on structure, organization, learning, traits, the self, character, attitudes, expressive movements, and types. Section III, *Dynamics of Personality*, talks about the biology of motives, briefly presents Freud, Adler, Lewin, and Allport, and works up to the author's synthesizing chapter on "A Cultural Interpretation of Motivation." Section IV, *Determinants of Personality*, presents material on biological factors, the family, the school system, economic factors, and on the relation between personality and social values.

The new book includes expanded treatment of several specific topics that have recently been receiving more attention. For example, the problems of reliability and validity are dealt with much more adequately than in the 1937 book, and there are now seven pages on "projective techniques" instead of one page on "the ink-blot test." But more important than the inclusion or expansion of specific topics are changes in general approach. For one thing, the author has taken into his thinking the concept of *frame of reference*, and with it, an increased attention to perceptual phenomena in personality. He has also stopped worrying as to whether personality is stimulus or is response, and has come to think quite comfortably in terms of intervening variables.

These changes in content and approach mean that the new book, more than the old, contains the raw material out of which the author's intended "adequate picture of the field" can be made. Whether the author paints an adequate picture may be open to debate. Those psychologists who are down with a rigid adherence to one or another conceptual approach to personality will not like the book's eclecticism. Others will be unhappy with the way in which the conceptual ingredients of the picture are arranged and presented.

The present reviewer, through the first 275 pages of the book, was bothered and puzzled by what appeared to be a tendency on the part of the author to enumerate rather than to relate various concepts—a tendency which will probably confuse all students and many instructors. For example, in dealing with the concept of *trait*, the author writes, "We need not necessarily consider the interpretation of traits as frames of reference to exclude either habit-systems or the mental-set interpretation. Each in its own way may help the student to grasp the functional character of traits" (p. 146). Such a *laissez-faire* approach to overlapping or closely related concepts tends to be repeated in the case of attitude, stereotype, and frame of reference, and to some extent in the early dealings with drive, valence and tension. Things seem to come out all right in the end, but the ride over the conceptual road is jerkier than one would wish—either for himself or for his students.

It is not a pretentious book. It is a book around which a good and comprehensive course in personality can be built. It is this reviewer's guess that students' reaction to the book will be generally favorable and that they will find the chapter on the social determiners of personality particularly provocative. By and large, the book seems to do what it sets out to do—to give the student "an adequate picture of the field, not a tightly organized theoretical interpretation" (p. vii).

FILLMORE H. SANFORD.

Haverford College.

SHELDON, WILLIAM H. (with the collaboration of E. M. HARTL AND E. MCDERMOTT). *Varieties of delinquent youth.* New York: Harper, 1949. Pp. xvii+899. $8.00.

This volume, the third in a projected series on Human Constitution, continues Sheldon's well known work on constitutional psychology. The first section of the book consists of an "orientation for a biological humanics." Next follows a section of some 600 pages containing abbreviated case histories complete with the somatotyping pictures and measurements of 200 delinquent youths studied at the Hayden Goodwill Inn in Boston. The final section discusses the psychiatry of delinquency. The book ends with a three-page epilogue which calls for a sweeping reorientation in our economic, political, sexual, and religious thinking, since "the problems presented by delinquency are inseparable from the underlying defections of social institutions."

The student of Sheldon's morphological doctrines will find here an application of his somatotyping technique to 200 subjects, a further development of the psychiatric implications of his constitutional approach, and an attempted application of them to the problems of delinquency. He will also find many stimulating and suggestive comments, although they seldom are buttressed either by original scientific data or by authoritative citation from the literature. To reach them he will have to wade through many pages of somewhat ill-tempered, authoritarian pronouncement which often ranges far beyond the topic presumably being discussed.

One feels that Sheldon has achieved not a careful scientific handling of an important topic, but rather a personal philosophical catharsis. The style is marked by a petulant omniscience that is irritating at first but finally ends by becoming unintentionally amusing. This reviewer deems it unfortunate that the author did not discipline himself to his scientific task. The publication of this book will do little to further an important and suggestive approach to human behavior, an approach already under heavy and sometimes undeserved critical attack, and ill-prepared to withstand the further criticism that Sheldon's performance will inevitably arouse.

WILLIAM A. HUNT.

Northwestern University.

NOTICE

The Editor's address is now *Department of Psychology, University of Illinois, Urbana, Illinois.*

Vol. 47, No. 6 November, 1950

Psychological Bulletin

RECENT RESEARCH ON PAIN PERCEPTION[1]

WARD EDWARDS
Harvard University

This paper is an outline of four related problems: the definition of pain, the psychophysics of pain, the peripheral nervous mechanisms involved in pain perception, and the role of the central nervous system in certain abnormalities of pain perception.[2]

I. DEFINITION

The literature on pain has few discussions of the problem of its definition. The prevailing view is outlined by Lewis in the introduction to his book, *Pain* (48):

> Pain, like similar subjective things, is known to us by experience and described by illustration. . . .
>
> We have no knowledge of pain beyond that derived from human experi-

---

[1] It is a pleasure to thank Dr. J. G. Beebe-Center, who interested me in this problem and gave me valuable advice throughout the preparation of this review, and Dr. E. G. Boring, who gave the manuscript a careful and critical reading. Their good advice led to the correction of many errors; those which remain must be attributed to my shortcomings.

[2] Several reviews of work on pain are available in the psychological literature, but none of them is very recent and none is exhaustive. Morgan's chapter on the subject in his *Physiological Psychology* (55) is the most recent review (1943) and deals mostly with the problem of pain receptors. Dallenbach (14), writing in 1939, covered thoroughly the history of the distinction between pain and unpleasantness, as well as dealing more briefly with other topics. Stone and Jenkins (74), in 1940, covered experimental work on pain and temperature during the 1930's. In the medical and physiological literature, several more recent reviews are available. An excellent review of the skin senses in general was published by Bishop (7) in 1946. In 1947, Wolff and Hardy (92) published another very good review. Wolff and Wolf (98) covered much the same ground in a short book in 1948. Wolff's earlier Harvey Lectures (89) are by now out of date, since they were published in 1944. Lewis's (48) book is a review of his own experiments for the most part (1942). Riddoch (63) reviewed the literature on phantom limbs in 1941, and Walshe (82) reviewed epicritic and protopathic sensibility in 1942. This paper was written to bring the reviews up to date, to make available in the psychological literature the results of medical and physiological research on pain insofar as they bear on psychological issues, and to propose some criticisms and some integrations which are not to be found in the medical literature.

> ence; yet we may judge of its presence in animals by bodily reactions that human experience has brought us to recognize as its frequent accompaniments or by the use of stimuli that similar experience tells us should be painful. . . . Strictly speaking, there are no reliable and usable indices of pain in animals; there are only phenomena recognized as frequent associations of pain. . . . Our knowledge of pain, of its sources, of the nerve paths conveying it, has been built up, and will continue to be built up, chiefly by observations upon man. . . . There is manifest wisdom in seeking knowledge from the direct, rather than from the indirect source, and here the direct source is prolific; there is, too, the deep satisfaction of knowing that any evidence won from man is indubitably applicable to the human problem (**48**, p. 1).

Many of the writers whose work is reviewed in this paper have stated views like Lewis's (e.g. **50**, **92**), and none of them has expressly disagreed with this essentially introspective definition of pain. This unanimity probably can be accounted for by the fact that the literature on pain is a medical literature. It seems likely that many behavioristically oriented American psychologists would prefer an objective, operational definition.

In the search for an operational definition of pain, the first step must be an examination of the criteria used by researchers in their work. By far the most commonly used criterion of pain is the verbal report (**29**, **48**, **53**). Occasionally, voluntary withdrawal of the stimulated area has been used as the criterion of pain in human beings (**12**, **67**). In animals, the most common criteria are cries, squeaks, or squeals; reflex responses of the area stimulated; struggling; and flight (**24**). Not always, however, do researchers rely on behavioral evidence that their stimuli are painful—or at any rate excite afferent pain fibers. It is possible to apply stimuli to nerves known on other evidence to carry pain fibers and to call these stimuli painful, whether or not they produce reports of pain. Or it is possible to record pain impulses directly from nerves carrying pain fibers (**66**) and to use these as the criterion of pain, or to apply stimulation to areas considered sensitive only to pain (**78**). Or a certain stimulus could arbitrarily be defined as painful on the basis of some other property it has—for instance, that it produces tissue damage, or that it is negatively reinforcing.

All the above may be *sufficient* conditions for use of the word *pain*. But a definition of a word consisting of a list of different situations sufficient for its use is not ideal. It would be much more desirable to be able to state a *necessary* condition for use of the word. If *pain* refers to something specific (as common sense urges that it does) then that necessary condition must be presence of the thing to which the word refers. What is that thing? Common sense says, the experience of pain.

If we continue to use the common-sense notion of pain, it is clear that no one of the above experimental criteria is adequate as a necessary criterion for use of the word. All of them can be shown to be absent in cases in which pain is present. Even excitation of peripheral pain fibers is absent in the case of "central pain" (**25**, **53**). The alternatives, then, are to decide that pain is not inevitably connected with any of its behavioral criteria or else to abandon the common-sense notion of pain. The medical literature, and therefore this review, prefers the former choice.

The basic fact here is that common sense, and therefore the pain literature, is dualistic in its epistemology. Pain experience is different from pain-produced behavior, and it is to the former that *pain* refers. This is an instance of a general problem of epistemology which has also arisen in interpreting research on other senses; it becomes particularly acute in pain research only because the stimulus for pain is so much less clearly understood than is that for other kinds of sensory experience. In other sense modalities, a particular kind of stimulation can be considered the criterion for use of a word referring to experience (e.g., light stimulation applied to an organism causes us to speak of color vision), and the few exceptions in which the experience occurs without the stimulus can be ignored, especially since they do not usually occur under the conditions of most experiments. In pain, no such simple dependence on an easily identifiable stimulus exists. For this reason, no operational definition of pain can be proved adequate until some generalizing operation can be found which will indicate identity among different cases of pain perception, and so justify in operational terms the use of the same word for these different cases. Whether such operations exist for other sense modalities is a debated question.

If pain is subjective, how can it be experimented on? Stimuli, responses, and experimental operations are used as indicators whose meaning may vary depending on the total situation in which they are observed. A wide variety of observations are made and interpreted by each experimenter in terms of his own common sense, generated from his past experience of pain, its stimuli, and its consequences. This procedure is subject to a variety of interpretations consistent with a variety of epistemological biases. But the basic requirements for any research are reproducible results and meaningful conclusions, and so long as these methods meet these requirements they are, in the opinion of this reviewer, entirely adequate.

The day will presumably come when functional relations involving indicators can be integrated by means of a neural model without any reference to the subjective concept of pain. Will this lead to dropping the subjective concept of pain from psychology? Perhaps. The history of vision, however, suggests that the answer is no. Despite advanced knowledge of physiological mechanisms in that field the concept of visual sensation is still the principal basis of organization for visual psychophysics.

II. DISTINCTIONS

A number of terms, most of them referring to dichotomies, have been used in the pain literature. A summary of these terms before discussion of the literature will provide a convenient entering place and a necessary vocabulary.

Pain-Unpleasantness. The distinction between pain and unpleasantness is fundamental to present-day thinking on the subject of pain. The word pain is now used to refer to a perception, like a tone or a color, rather than an affective state or a performance in a choice situation. It is in this sense that the word is used throughout this review. The history of the emergence of this viewpoint was reviewed by Dallenbach (**14**) in 1939.

Pain Perception–Pain Reaction. Recent workers in the field, particularly Hardy, Wolff, and Goodell, have emphasized the distinction between pain as perceived and the behavior which may or may not result from that perception. This distinction is favored by experiments showing that pain reaction threshold, as measured by GSR, may vary independently of pain perception threshold, as measured by verbal report, under the influence of analgesics or spontaneously (**94**, **97**; see also **12**, **67**, **83**).

Skin Pain–Visceral Pain. Within the field of pain perception, it is usual to distinguish between pain coming from the receptors in the skin and pain coming from muscles or viscera. Most experimentation has been done on the former, since the skin is more accessible to easily controlled stimuli. However, clinical experience suggests that visceral pain is the more important of the two in the life history of most persons. This distinction as to origin does not necessarily imply a similar distinction as to quality; whether visceral pain is qualitatively the same as skin pain is a disputed question (**18**).

First Pain–Second Pain. When the skin is pricked, there will sometimes be a distinct double experience of pain. The first pain is usually described as sharp and pricking; the second is a longer, more burning pain. This qualitative distinction between first and second pain is also in dispute. In general, only the extremities show this double pain.

Pricking Pain–Burning Pain. The distinction between pricking and burning pain is more general than the first pain-second pain distinction. Pricking pain is, in general, reported for intact skin, while skin in which nerves are regenerating, which has been partially denervated, or which has been injured so as to cause hyperalgesia yields burning pain on appropriate stimulation. Visceral pains have been called similar to burning pain. It is a matter of controversy whether or not these are differences in anything more than length of the experience (2, 48, 87).

Epicritic Sensibility–Protopathic Sensibility. Henry Head (36) in 1905 used these terms to describe the results of his experimentation on cutaneous nerve regeneration. Epicritic sensibility, the ordinary sensibility of the skin, is accurately localized and not painful. Protopathic sensibility is diffuse and unreasonably painful, has a higher threshold than epicritic to most forms of stimulation, and does not give experiences of touch. His theory was that these two forms of sensibility are mediated by two forms of innervation. The protopathic nerves are more primitive and regenerate more quickly than the epicritic, but are ordinarily inhibited by the epicritic nerves in the central nervous system. Nerve division at the periphery, the spinal cord, or the thalamus will stop the inhibitory function of the epicritic system in the areas affected, leaving the unpleasant protopathic system in action, if there is any effective innervation of the affected region at all. It is a question whether Head would have equated this distinction, so far as its reference to pain is concerned, with the distinction between pricking and burning pain discussed above. Most experimenters have failed to repeat Head's observations (9, 79, 82), so the concepts of epicritic and protopathic sensibility have fallen into disuse.

Neuralgia-Neuritis–Hyperalgesia-Causalgia. Neuralgia is a constant pain in the distribution of a particular sensory nerve, which may or may not be caused by neuritis, which is an inflammation of a nerve. Hyperalgesia is an increased sensitivity to painful stimuli in a restricted area. Causalgia is a syndrome which includes hyperalgesia and a constant burning pain. It is usually the result of a nerve lesion.

III. Pyschophysics of Pain

The subject of pain has been the province of the phenomenologist, the physiologist, and the physician, especially the latter. It is a difficult area in which to do experiments. Stimuli which will produce only pain are difficult or impossible to come by; the quantitative relationships between the stimulus as controlled by the experimenter and the effective stimulus at the receptor are unknown; and the range in which painful stimuli can be applied without causing excessive tissue damage is small. Subjects willing to expose themselves to painful stimuli may be hard to find. The responses used as indicative of pain are not always highly correlated, and any strong painful stimulus is likely to generate an emotional state which makes detached observation and report of the resulting pain difficult.

The literature to be reviewed here is mostly a medical literature. This means that it includes a good deal of case material and that the experimental studies frequently are inadequately controlled. Very few of the studies reviewed here reported standard deviations; even fewer made tests of significance. Clearly the quantification of our knowledge of pain has hardly begun. However, this should not be taken as adequate reason for discounting the value of the conclusions which medical experimenters or clinicians have drawn. The qualitative problems of pain are numerous and difficult, and in some of them the informed opinion of a doctor who has been treating patients in pain most of his life may be more useful than any experimental attack could be. In any case, until more carefully controlled studies are done and better methods of attacking the problems of pain are available, the information from these studies remains the best there is in the field.

In the last decade a wholehearted attack on the problem of quantifying the field of pain in a manner analogous to the psychophysics of vision or hearing has been made by J. D. Hardy, H. G. Wolff, H. Goodell, and their associates at the New York Hospital–Cornell University Medical College. This whole program has been based on the experimental technique of measuring pain thresholds or other quantities by means of a projected beam of heat and light which stimulates a blackened area of the skin (29). The exposure was for three seconds (as a rule), and the intensity of the beam was controlled by a rheostat. In later experiments, a variac was used. Stimulus intensity was measured with a radiometer, or, in later experiments, with a scaled voltmeter calibrated to give equivalent measures. Subjects (usually the experimenters) were instructed that they would experience "warmth mounting into heat which would suddenly, just at the end of a given exposure, 'draw together' into a sharp prick or tingle (2)." The threshold was approached from below only, using a variation of the method of limits. The point at which the subject reported pain was taken as his threshold.

There are difficulties in this procedure, of course. It does not provide for the statistical manipulation of data which has been found necessary in the psychophysics of other sense modalities. The approach only from below is likely to introduce a constant error into threshold readings. The use of a radiometer reading as an index of the amount of thermal energy delivered at the skin depends on the thoroughness and constancy from person to person of the blacking of the skin. Most important, Lewis (48) has shown that radiation applied to the skin will produce hyperalgesia, due to the release of histamine (65) or a histamine-like chemical (48) into the skin. Schumacher (68), using the Hardy-Wolff-Goodell apparatus, has shown the same thing for ultraviolet radiation. It is possible that the weak stimuli used in threshold determination do not have this effect; but if they do, the measures which Hardy, Wolff, and Goodell have made must have been seriously affected. The earlier papers are not explicit as to where successive measurements were made, but every description given suggests that successive measurements were made at the same place on the skin. It is only in a paper published in 1949 (27) that the first warning appears not to present successive stimuli to the same area of skin. (This warning is in connection with suprathreshold stimuli, but previous papers [32, 33] also used suprathreshold intensities, without warning. In these previous papers it was recommended that exposures be spaced out: five stimuli in ten minutes in the first of them, and five minutes between stimuli in the other.) Finally, none of the papers reports tests of significance to justify conclusions, and few of them report variability data in such a form as to make such tests possible. All these experimental difficulties should, of course, be corrected if possible in future experimentation, but the results reported by Hardy, Wolff, and Goodell are apparently fairly stable and reproducible (12, 73) and it seems unlikely that their more important conclusions will be changed too much by future studies.

In their original paper, published in 1940, Hardy, Wolff, and Goodell (29) reported that the pain threshold for their own skins was approximately 0.23 g. cal. per sec. per sq. cm. of skin surface stimulated. Pain sensitivity was separable from touch, heat, and cold, in that stopping circulation in an arm destroyed touch, heat, and cold sensitivity after 25 min., while the pain threshold, though raised, was still measurable. Furthermore, it was demonstrated that pain elsewhere could raise the pain threshold as much as 35 per cent. The pain generated by stoppage of circulation of the arm was found to raise the pain threshold at the forehead by the same amount as the threshold at the arm was raised, suggesting that this rise in arm threshold was not due to the ischemia produced by circulation stoppage. Also, acetylsalicylic acid (aspirin) was shown to lower the warmth threshold and raise the pain threshold, indicating that the two are separate senses, although stimulated together in this technique. The size of the area of skin stimulated was shown not to influence the amount of the threshold, indicating that there is no spatial summation for pain (in contrast to heat, for which spatial summation is almost perfect [28]), and justifying the authors in reporting thresholds in terms of a physical unit independent of the amount of skin stimulated.[3]

[3] Gram calories per second per square centimeter of skin stimulated, the unit used, is actually independent of the amount of skin stimulated. As the amount of skin stimulated is increased, the intensity of the light remaining the same, the total energy in gram calories delivered to the skin will increase, since area increases, but the energy per unit of area per unit of time will remain the same.

Following this basic paper, Wolff, Hardy, and Goodell reported a series of studies on the effects of various analgesics on the pain threshold (**93**, **94**, **95**, **96**, **97**). The principal findings were:

1. Most ordinary analgesics raise the pain threshold (**94**, **95**).
2. The extent of the rise depends on the amount given, but there is a maximum rise which any given analgesic can produce. This maximum rise may be as much as 100 per cent (**94**, **95**).
3. Intake of quantities greater than adequate to produce maximum rise will in general prolong the effect of the analgesic (**94**, **95**).
4. The pain-threshold-raising properties of analgesics given in combination do not summate; such a combination is only as effective as its most effective ingredient (**94**, **95**).
5. The opiates (**94**), ethyl alcohol (**97**), and to a lesser extent other analgesics have the additional effect of dissociating pain perception from pain reaction. Subjects may perceive considerable pain, but do not react to it. This effect does summate and has no observed limit.
6. If serious pain has existed for a while before administration of an opiate, such administration will not affect the pain threshold, but will affect the pain reaction (**94**).
7. It is concluded that the reduction or removal of the pain reaction is the main value of opiates, and is the most important effect for which doctors administering analgesics should try (**94**).

The results of these studies, however, are called into question by another study by Wolff and Goodell (**91**), in which subjects were administered placebos. The results were that, in a suggestible subject, a placebo believed to be an analgesic will raise the pain threshold while an analgesic believed to be a placebo will have no effect on the pain threshold. (The latter conclusion was demonstrated only for acetylsalicylic acid; no other analgesics were tried.) No data were given as to whether or not the subjects of the previous experiments on real analgesics were suggestible.

So far, all the studies described have exposed the stimulus for three seconds and instructed the subjects to report the threshold of pricking pain. Bigelow, Harrison, Goodell and Wolff (**2**) tried to compare the thresholds for pricking and burning pain under various conditions. They found that the threshold for burning pain was lower than that for pricking pain. There are, however, two experimental difficulties. First, the differences as graphed are so small that a test of significance is needed to determine whether or not they are random, particularly since the data for only eleven of the twenty-three subjects used are included in the graph from which the conclusion cited was drawn. Furthermore, the authors reported that it was easier to judge the pricking pain threshold at the shorter exposure lengths, while longer exposure lengths were best for the burning pain threshold. This leaves some doubt as to the validity of reports of both measures for the same exposure time. This doubt is strengthened by Lewis's (47) contention that actually the only difference between pricking and burning pain is the length of stimulation, a view which Bigelow, Harrison, Goodell, and Wolff do not share. It is not agreed whether the first of the two cutaneous pain responses from an extremity is pricking and the second burning pain, so Bigelow, Harrison, Goodell, and Wolff tried to settle the issue by asking nineteen subjects. The subjects disagreed: eight said the two pains were of the same quality, nine said they were different, and two were undecided. The authors also measured the two thresholds on the glans penis, which Head (36) had considered to have only protopathic sensibility. Here the two thresholds were widely different, with the pricking pain higher. The introspective reports, however, leave much doubt as to whether the subjects were judging the same kind of pricking pain in this case as they were in measurements on the arm or forehead, as the authors point out. The authors report several other experiments involving ischemia, which abolishes the first pain, or procaine nerve block, which abolishes the second. In all of these, the first pain was felt as pricking and the second as burning. However, the duration of pain was uncontrolled for all of these cases.

Hardy, Wolff, and Goodell (31) proceeded in 1947 to the next step in the psychophysics of pain—the measurement of the just noticeable difference. They found that the ratio of stimulus intensity at threshold to stimulus intensity adequate to produce tissue damage was only 1:2, which they interpreted in terms of the biological inefficiency of a pain sensitivity which would extend much farther below the point of tissue damage. Within this range, they reported that there were 21 j.n.d.'s. The size of the j.n.d. changed, but was approximately 0.01 g. cal. per sec. per sq. cm. in the lower ranges of stimulation. The method used was again a variation of the method of limits, open to the same criticisms discussed above.

Now the ground was clear for the construction of a scale of pain intensity, which the next paper (32) did. This paper determined the j.n.d. at various intensities from threshold to the point at which changes in intensity could no longer be discriminated, and then used these j.n.d.'s to compute a double j.n.d. scale. The units of the scale were named *dols*, and the range of the scale was, of course, $10\frac{1}{2}$ dols (corresponding to 21 j.n.d.'s). The values of the j.n.d. used in calculating the scale were, however, adjusted somewhat from the original data. How much difference this adjustment makes is not clear.

A follow-up study (33) attempted to valiate this scale by the method of fractionation. The validation was carried out both on the authors and on seventy medical students. The authors assert that the validation is close, but failure to report variability data or to test significance makes it impossible to evaluate this conclusion. The average figures presented look like good validation.

Finally, Hardy and Javert (27) tested the usefulness of the dol scale by measuring the pain intensity of childbirth, using estimates of absolute intensity in dols. The only theoretically important conclusion was that once again the distinction between pain perception and pain reaction was necessary to make sense out of the results. The maximum pain perceived ranged from about 7 dols (without anesthesia) to $10\frac{1}{2}$ dols, the top of the scale.

This is the most recent in a long series of related experiments by the same group, using the same apparatus, and directed towards the same end: the quantitative examination of pain perception. The series, taken as a whole, seems to offer three conclusions:

1. It is both possible and desirable to work out the psychophysics of pain.
2. A distinction between pain perception and pain reaction is probably necessary in order to organize the experimental results in any meaningful way.
3. As the psychophysics of pain progresses, it will become increasingly necessary for experimenters to keep pace with the quantitativeness of their problems by refinement of their experimental and statistical procedures.

A number of other techniques have been used, with varying degrees of success, in the attempt to measure pain thresholds (8, 24, 35, 73). Slaughter and Wright (73) have modified the Hardy-Wolff-Goodell apparatus and procedure by the addition of a constant voltage regulator to avoid fluctuations in heat delivered due to fluctuations in line voltage, the use of a small thermocouple just in front of the subject's skin to measure strength of stimulus at the time the stimulus is applied (in contrast to the later use of a radiometer by Hardy, Wolff, and Goodell) and the reporting of results in degrees. The first two improvements seem desirable. The third is open to objection. The effective stimulus to pain in a situation in which radiant heat is applied to the skin is presumably heat transfer. Heat transfer is dependent on the amount of heat delivered to the skin, but only secondarily dependent on the temperature at which the heat is delivered. Actually, the two measurements should be convertible one into the other, since both are made before the stimulus reaches the subject. Using this technique, Slaughter and Wright (73) have found much more inter-person variation in pain threshold than did Hardy, Wolff, and Goodell. Leduc and Slaughter (46) have in general confirmed the Hardy-Wolff-Goodell results with opiates.

Goetzl, Burrill, and Ivy (24), after an extensive review of the algesimetric literature, proposed (in 1943) another method which they considered more desirable than any proposed up to that time. This method consisted of applying pulses of alternating current to a tooth. The intensity and frequency of occurrence of the pulses could be varied. The procedure was used on men and on dogs. On men, the current was applied through a platinum electrode into a filling, and the other electrode was a bath of salt water into which the hand was put. On dogs, two holes were drilled, one on each side of a tooth, and the holes were filled with fillings especially hollowed out to receive the electrodes. In this case no indifferent electrode was used; the current was passed from one side of the tooth to the other. A reflex contraction of the facial muscles was considered the sign of pain. Using this technique, Ivy, Goetzl, Harris, and Burrill (**38**) showed that epinephrine has analgesic properties, and Parsons and Goetzl (**57**) and Goetzl (**23**) tested the effects of acetylsalicylic acid and other antipyretic drugs on the pain threshold. They concluded that antipyretic drugs do not have pain-threshold-raising properties. Parsons and Goetzl (**56**) also confirmed the observation by Hardy, Wolff, and Goodell (**29**) that the presence of one pain raises the threshold for another. Robertson, Goodell, and Wolff (**64**) later used a modification of Goetzl, Burrill, and Ivy's technique and found considerable variation in threshold value observed depending on placement of electrodes, recent stimulation, etc. It is not surprising that such a technique yields variable results; the resistance of the tooth to the passage of current is certainly likely to be important in determining the amount of current necessary to produce a report of pain; this resistance is very likely to vary from subject to subject, and maybe even from time to time for one subject. On the other hand, this technique is obviously superior to the Hardy-Wolff-Goodell technique for animal experiments, since it requires no particular cooperation from the animal. It is probably the best technique yet available for pain experiments on animals.

Lewis (**48**) has proposed and Harrison and Bigelow (**35**) have systematically used a technique for measuring pain threshold in ischemic arm muscle. The circulation of the arm is blocked, and the subject makes a series of hand contractions of constant size, frequency, or energy. The number of contractions needed to elicit pain is used as the threshold figure. It is necessary, of course, to allow enough time between successive threshold determinations for dissipation of the chemical which causes ischemic pain (**48**), and to avoid muscle fatigue, which raises the threshold (**35**). With these precautions, this method yields highly constant thresholds, which are amenable to experimental alteration in a manner analogous to the experiments of Hardy, Wolff, and Goodell and which are not subject to the difficulties which repeated exposures to radiant heat present. The main advantage of this technique, however, is that it permits a better determination of the end-point for pain. So long as the muscle is not used, whatever degree of pain has been developed in it remains fairly constant. This makes possible relatively leisurely observation of the experience and so presumably an easier judgment as to whether or not pain is actually present.

The experiment of Denton and Beecher (**15**) is unique in the pain literature for its statistical sophistication (the authors give credit to a mathematical statistician, Dr. F. C. Mosteller, for statistical advice). This is an elaborate experiment attempting to compare the analgesic and other effects of several drugs. In most respects it is a model of experimental design. The pain used, however, was pain resulting from many different kinds of surgery, and the comparison of analgesic effects naturally becomes less convincing because of the variation in amount and kind of pain which the analgesic was used to control. The experimenters tried to randomize administration of each analgesic for these variables, but it is impossible to tell whether this was successful or not. The fact that 429 patients were used, randomized for type of operation over three types, suggests that they may have been successful. In this experiment, no pain thresholds were measured; all measures were of the amount of analgesic required to give relief from pain to 90 per cent of patients receiving it. A parallel experiment performed on healthy subjects was designed to determine the side-effects of the drugs. Among other procedures, the Hardy-Wolff-Goodell technique was used, but the results were so unsatisfactory that they were not even reported. For example, a salt solution frequently showed as much analgesic effect as morphine, or more.

IV. Skin Pain Receptors and Afferent Neurons

Writers on the receptors and afferent nerves involved in pain have centered on three main problems. First, there is the old von Frey-Goldscheider controversy as to whether pain has separate receptors. This issue has been almost completely settled in favor of von Frey and separate receptors. (For a discussion of the evidence leading to this conclusion, see Morgan [**55**].) If more evidence were needed, the dissociation of pain thresholds from those of other kinds of cutaneous sensibility demonstrated by Hardy, Wolff, and Goodell (discussed above) would serve. Not everyone is satisfied, however; Hebb (**37**) has argued very recently for a return to the intense-stimulation theory. His views will be discussed later in connection with the problem of brain functioning in pain.

The second problem with regard to skin pain receptors is raised by the double pain experience mentioned above. In general, this double pain from a single stimulus can be excited only on the extremities, which has led to the theory that it represents the different conduction times of different pain nerves en route to the central nervous system (10, 22, 48). Lewis (48) and Gasser (22) have each summarized the evidence for this interpretation, which is mostly derived from the results of different techniques of anesthetization of nerve fibers. Thus, Gasser (22) has shown that cocainization abolishes first the response of slow unmyelinated fibers. Correspondingly, Lewis (48) has shown that cocainization either of skin or of the afferent nerve first abolishes the second pain response. Gasser also demonstrated that ischemia abolishes the response of fast fibers, and Lewis (48) and Zotterman (101) have shown that it abolishes the first pain response. However, the evidence is not as yet complete; the quantitative temporal relations of ischemia data conflict as between Lewis and Zotterman, and no one has combined ischemia or cocainization with recording of the nerve impulses. The technique of Rusinov (66) for recording pain impulses should be adaptable to this problem.

The third issue to be considered here is that of the nature, distribution, and functioning of the end organ for pain. Here, too, the evidence is pretty clear. Four different histological investigations of various skin areas have been made, and all agree as to the nature of the end organ for pain (6, 78, 85, 99). Tower (78) examined histologically the cornea of the cat (the cornea is believed to be sensitive only to pain) and found "plexiform arrangements of nerve fibers . . . , bearing free nerve terminals in the form of simple terminal twigs, end knobs, loops, brushes, or minute skeins. . . . The terminal portions of this ramification are in all sites unmyelinated, but farther back, myelinated or unmyelinated fibers may enter into it (p. 17)." Such structures are found throughout the cutaneous surfaces, both within and beneath the epidermis. In general, the epidermal structures feed into fast, myelinated fibers, while the subepidermal ones feed into slower fibers.

The most important part of this conclusion is that such structures overlap and are widely ramified; in the cornea of the cat one fiber may receive impulses from stimulation of as much as one quarter of the whole cornea (78). This raises a serious problem: how is localization of pain as accurate as it is, and how do we get the so-called pain spots? Long ago, the pain spots were recognized as varying, and this picture suggests the view that they merely represent accidental variation in the density of receptors stimulated. The problem of localization is more difficult, but the anatomical evidence forces the conclusion proposed by Tower (78) that the pattern of stimulation in overlapping pain receptor units feeding into different nerve fibers is "interpreted" by the brain in terms of localization. As Wolff and Hardy (92) put it: "The painful spot experienced as such when an area of normal skin is exposed to noxious stimulation is a centrally integrated experience projected onto the periphery for the specific purpose of localization. The 'skin spot' is, so to speak, a 'mind spot' (p. 168)."

Evidence in favor of this view, which has the added advantage of settling the problem of sensation during regeneration after nerve division, is offered by Weddell, Sinclair, and Feindel (86). These investigators made histological examination of partially denervated skin after degeneration due to nerve section. They found that when the area of skin stimulated had isolated nerve nets, the characteristic slow, excessively unpleasant, poorly localized pain observed during nerve regeneration was elicited, but that when the area had many interweaving nerve nets and terminals, as in normal human skin, normal pain perception was elicited. This experiment accounts for all of the facts of nerve regeneration except the original ring of hyperalgesic tissue at the periphery of the dead area immediately after the section of the nerve. This ring of tisue is accounted for by the observation by Weddell (85) of overlapping innervation of any one skin area by many fibers from different nerves. If the fibers from one nerve degenerate, this leaves the fibers from the other nerves innervating the same area of skin alone, which is the condition prescribed by Weddell, Sinclair, and Feindel for the hyperalgesic response.

The problem of the end organ for pain, and the related problem of whether or not there are different qualities of pain, is neatly rounded off by an observation by Feindel, Weddell, and Sinclair (18). They made histological studies of various deep pain-sensitive tissues, and found in every case the same pattern of isolated nerve nets found in partially denervated skin tissue. This accounts both for the relative insensibility of deep structures to ordinary stimuli and for the excessively unpleasant nature of visceral and deep pain, when aroused.

The research discussed in this section, taken all together, justifies the conclusion that there is fundamentally only one kind of pain receptor, which gives ordinary pain sensibility when enough receptors are stimulated, but the kind of peculiarly unpleasant pain found in partially denervated tissues when too few are stimulated. It is a high density of these receptors, discharging into different nerves, which accounts for a "pain spot."

V. DISORDERS OF PAIN PERCEPTION

Aside from the problems of psychophysics and end-organ physiology, knowledge about pain is spotty and confusing. The observations of Weddell, Sinclair, and Feindel, cited above, suggest that much of what is important in the perception of pain goes on in the central nervous system. All that is known about disorders of pain perception heightens this belief. The history of experimentation in vision has been a history of experimentation on the peripheral mechanisms giving way to experimentation on brain mechanisms, as it became clearer and clearer that what happened in the eye was only a part of the whole problem of vision. In hearing, the problem of the role of the central nervous system is still more acute, but the experimentation has not for the most part been done. In pain perception (and, presumably, the skin senses in general) knowledge has hardly advanced beyond the point of asking questions about what happens in the central nervous system. The most important conclusion which can be drawn from the facts about pain and from the history of research in other sense modalities is that the picture will never be clear until the functions of the central nervous system in pain perception are well understood. The following discussion of abnormalities in pain perception illustrates this point.

The pathways which pain fibers follow are fairly clear as far as the thalamus. Skin pain fibers in general are conveyed to the dorsal roots by somatic nerves. Deep pain fibers may attach themselves to blood vessels for a while and then join autonomic nerves, may join autonomic nerves early and remain with them almost up to the cord, or may travel by somatic nerves. It does not seem to make any difference to the quality of the pain which path is followed. On entering the cord, all pain impulses are transferred to the opposite side and travel up the cord in the anterolateral section. From here, the path is uninterrupted for most pain impulses to the nucleus centralis posterior of the thalamus. Whether or not pain impulses are further projected to the post-central gyrus is uncertain; Penfield and Boldrey (**59**) were unable to elicit pain by stimulation of this area and hence concluded that pain is primarily represented in the thalamus, but analgesia and hyperalgesia have been reported in clinical cases involving cortical lesions (**53, 81**). Hemispherectomy does not result in any significant loss in pain sensibility on the affected side, but this operation apparently results in extensive thalamic degeneration also (**81**). The retention of sensibility suggests a pathway to the ipsilateral side of the cortex. Also, a few cases involving very extensive thalamic destruction still show pain sensibility. In short, brain localization of pain functions is still very much an open question.

Probably the most interesting and dramatic type of disorder of pain perception is known as causalgia. Livingston (50) has described a large number of cases of this remarkable syndrome, enough so that certain constant features stand out. The initiating factor may be an injury of as traumatic a nature as an amputated limb, but lesions as small as a thorn prick have been demonstrated to produce causalgia, and a number of cases are available where no lesion has been found. The symptoms are: incessant burning pain from the affected region, extreme hyperalgesia from the skin in that region, lowered temperature in that region, relatively small areas showing even greater sensitivity than the "normally" hypersensitive area, muscle spasm (in some cases), excessive sweating in the affected region, degenerative tissue changes, and marked personality changes.

A number of these symptoms reflect the hyperactivity of the sympathetic nervous system that seems to be a standard feature of all cases of intractable pain. Leriche (47) has based his whole conception of the treatment of intractable pain on the concept of sympathetic involvement, a view which seems more plausible when it is remembered that many afferent pain fibers pass through sympathetic tracts (3). Another important feature is the presence of muscle spasm. It has been suggested that in many cases of intractable pain the compression of nerve trunks by spasm of adjacent muscles (or viscera) is a causative factor (72).

More significant than the nature of the syndrome are the techniques effective in curing it. Occasionally cutting of the pain pathway involved at the peripheral or spinal levels will be adequate, but more often such a procedure will fail. The most dramatic technique is the injection of novocaine into a supersensitive point in skin or muscle or a sympathetic ganglion. This may relieve pain for a short while, or, more dramatically, it may relieve the pain permanently (26, 50). Such results, of course, suggest that some recurrent circuit mechanism was active, which once stopped did not start up again. Livingston's most dramatic case, making a very forceful argument for this interpretation, was of a man who developed causalgia of the left foot after a fracture. The middle three toes were flexed in spasm. After a few years, the corresponding three toes on the right foot began to flex, and simultaneously the right foot, which had not been injured, became slightly causalgic.

Amputation of the affected toes of both feet left very painful phantom toes on the left foot, which were eventually cured by amputation of that foot. Here the spread of the symptoms to the right foot after injury and prolonged pain in the left foot clearly suggests some active process in the central nervous system, which gradually recruited more and more of the fibers in that area into its activity.

A particular kind of causalgia which is very informative because it is so frequent is that resulting from limb amputation (63). Such amputations almost invariably leave phantom limbs, which feel to the patient just like real limbs. In non-painful cases the phantom limb can be moved and used, feeling just like a real limb. But in the very common painful phantom limbs, the phantom is usually tightly flexed, immobile and very painful. Such causalgias may be caused by neuromas in the stump, in which case excision of the neuroma may cure the pain. But frequently there is nothing wrong with the stump, and sectioning of the nerve trunks innervating it does not relieve the pain. Here again, anesthetization or destruction of sympathetic ganglia may result in permanent or temporary remittence from pain. It is interesting that along with the remittence from pain the patient regains the ability to manipulate his phantom limb. A frequent observation is that a resection of a nerve serving the stump may yield remittence for a while, with subsequent return of the pain. Whether this is due to growth of adjacent nerves into the denervated area or whether it is due to perpetuation of the central pain process has not been determined.

That these results are not due to psychogenic factors is suggested by Livingston's observation that injection of saline salution into a hypersensitive "trigger point" or a sympathetic ganglion has no effect, while injection of novocaine may have the dramatic effects described above. Livingston (50), from whose treatment and Riddoch's (63) the above picture was drawn, interprets the phenomena of causalgia as follows:

> An organic lesion at the periphery, involving sensory nerve filaments, may become a source of chronic irritation. Afferent impulses from this "trigger point" eventually create an abnormal state of activity in the internuncial neuron centers of the spinal cord gray matter. The internuncial disturbance in turn is reflected in an abnormal motor response from both the lateral and anterior horn neurons of one or more segments of the cord. The muscle spasm, vasomotor changes, and other effects which this central perturbation of function brings about in the peripheral tissues, may furnish new sources for pain and new reflexes. A vicious circle of activity is created. If the trigger point is removed early, the process may subside spontaneously. If the process is permitted to continue, it spreads to involve new areas, and tends to acquire a momentum that is increasingly difficult to displace. Perhaps in this stage even

> a removal of the original irritant may not be sufficient to establish a cure. But if an important part of the circle of reflexes can be interrupted, the process may subside, and a normal physiologic status is again established (50, pp. 226–227).

This view is a far cry from a simple one-receptor–one-sensation view of pain. And it is not at all unique to investigators in this field. Russian investigators since the war, operating on a similar hypothesis, have found that up to one hundred hours of continuous sleep is an excellent treatment of causalgia (though they were working only on newly injured cases [39, 100]). They have also had good results with novocaine block of the affected nerve (80). Cohen (13) has produced the pain of angina by application of an irritating plaster to skin innervated by the same segment as the heart in patients with heart disease, a result which he interprepts in terms of summation of subliminal pain impulses in the central nervous system. Bishop (4), from his experiments on electrical stimulation of the skin, concludes:

> Even though electrical stimulation of one ending of one fiber may be possible, it does not follow that the central nervous system transmits and interprets the impulses induced in a correspondingly simple manner. The chief virtue of such experimental restriction of activity to one peripheral element may be the demonstration that such stimulation is insufficient to enable the central nervous system to obtain an adequate estimate of external conditions. It may be inferred that the central nervous system is designed not primarily to detect discretely the activity of each ending, but rather to integrate the activities of many; and even peripheral sense organ activity will not be satisfactorily analysed, at least in terms of a subjective sensory response, without a further understanding of the central mechanisms that react to it, as well as a knowledge of what each sense organ does to its immediate nerve fiber (4, p. 379).

A variety of evidence about this notion of a spinal interaction among pain processes is available, most of it favorable. For example, Kunkle, Charles, Armistead, and Goodell (45) have shown that pain produced by immersion of one finger in cold water spread to other fingers, but that pain did not so spread from the thumb to the jaw (whose representations in the post-central gyrus of the cortex are adjacent), and so concluded that the neural process involved is in the spinal cord. Thompson and Gellhorn (77) have shown that painful stimuli can produce spinal reflexes in spinal animals. On the other hand, Bingham (3) argues that the effectiveness of anesthetization or destruction of sympathetic ganglia in curing causalgias is the result of blocking of sympathetic afferents and not of stopping the support of some ongoing spinal activity; and Doupe, Cullen, and Chance (16) believe that causalgia is caused by direct stimulation of pain afferents by sympathetic efferents.

Aside from the recruitment and interaction effects postulated above, there is also the concept of inhibition. The view of Head (36) was that epicritic fibers inhibited protopathic, and this view is still widely held in one form or another. The most common view is that the impulses resulting in fast pain modify the central nervous system in such a way as to prepare for the subsequent impulses resulting in slow pain (42).

The only detailed theory of the functioning of the central nervous system in pain has been offered by Hebb (37). His theory of pain is interwoven with the rest of his theory of brain function, but it can be summarized by saying that he conceives of pain as being a massive disorganized firing of thalamic or cortical neurons. This massive, disorganized firing may develop if too few impulses from C fibers reach the central nervous system, allowing localized hypersynchrony of firing to take place in the thalamus. It may also take place if impulses along C fibers arrive so frequently as to dominate local thalamic action, again producing localized hypersynchronous firing. In the intermediate normal range of action of C fibers, their effect is perceived as touch. To account for the fact that injury of the somesthetic system anywhere from peripheral nerve to thalamus may produce pain, and to account for the nerve division experiments of Head (36), Boring (9), and others, Hebb postulates that hypofunction of the somesthetic system leads to pain. He assumes that more general hypersynchronous firing, or more generally organized firing of any kind, will not be accompanied by the experience of pain; only local hypersynchrony is painful. The reason why local hypersynchrony may develop is that in the absence of the arrival of continuous normal somesthetic impulses, "Higher somesthetic nuclei then lack their normal degree of sensory control and become hyperresponsive, and yet are still exposed to an intermittent 'triggering' from peripheral stimulation" (37, p. 188).

The notion of hypofunction as a cause of pain is a new and imaginative one, but it raises the question why operations attempting to cure intractable pain by severing the pain pathways, either peripherally or in the spinal cord, are ever successful. Diseases such as syringomyelia, in which afferent pathways degenerate without leaving the patient in pain, are also unexplained by Hebb's theory, as he has pointed out. It is probable that the usefulness of Hebb's theory of pain will depend on the usefulness of his view of the function of the central nervous system.

It would be misleading to leave the topic of central nervous system functions in pain disorders without considering some of the disorders which are or may be peripheral in origin and cause. Of course, neuralgia due to neuritis is presumably peripheral in origin. Lewis (48) has performed a series of experiments showing that hyperalgesia of the skin such as is produced by sunburn is caused by the release of a chemical (histamine-like, for Lewis, but Rosenthal and Sonnenschein [**65**] argue that it actually is histamine) into the skin. Such hyperalgesia occurs in the area of distribution of the nerve innervating the area damaged, when it results from localized damage (e.g., that due to repeated strong electrical stimulation of a small area). Lewis blocked a nerve with novocaine and then stimulated the nerve trunk either above or below the block. If the stimulation was applied above the block, pain was felt, but no hyperalgesia resulted. If it was applied below the block, no pain was felt, but hyperalgesia resulted after the block wore off. Lewis concluded that the effect depends on the spread of nerve activity, and that what happens is that noxious stimulation causes the nerves stimulated to release the histamine-like chemical in the surrounding area. He attributed all this to a special set of nerves which he called the *nocifensors*, and which have no functions other than those described above. He postulated these nerves because the facts of localization seemed to him to require punctate distribution of pain receptors and neurons, which is not consistent with the diffuse distributions shown by the hyperalgesia. However, the more recent work on localization and pain spots described above disposes of this problem, and leaves it as probable that Lewis's results are caused by antidromic impulses in the sensory fibers of the pain system—an interpretation accepted by many of the more recent writers on the problem (**78, 87**). This looks like a clear demonstration of a peripheral mechanism for one kind of hyperalgesia, but it may not be. In 1950, Hardy, Wolff, and Goodell (**34**) published a series of observations on hyperalgesia which they feel must be explained in terms of an ongoing facilitative activity in the spinal cord, of much the same type as that described above by Livingston. They found, for example, that spatial summation of pain impulses occurs in hyperalgesic skin but not in normal skin. It remains to be seen whether the peripheral theory of hyperalgesia can be extended to cover their observations.

Another phenomenon which unfortunately has not been adequately examined, is the cases of people who are congenitally insensitive to pain. That such cases exist in very clear (**19, 44**). No histological studies of the skins of such people have been reported. But the presence of spinal reflexes which are assumed to be initiated by pain impulses suggests that the defect is in the central nervous system.

A good deal of research has been done on the mechanisms underlying various other kinds of pains, especially visceral pains (**35, 52, 87**) and headaches (**51, 70, 71, 88, 90**). This work will not be reviewed here, since it has contributed nothing new to the psychology of pain in general, aside from a demonstration of the great frequency of inaccurately localized or "referred" visceral pains. Discussions are also available of the various surgical techniques for curing intractable pain (20, 47, 76) including particularly cutting of the posterior roots (60) or of the anterolateral tract (75). Recently a number of surgeons have begun to perform prefrontal lobotomies to cure intractable pain (17, 83). One interesting result of this is that a patient who formerly was completely dominated by his pain will after operation show no pain symptoms and be able to go about his business with no impairment of function due to pain; but if he is asked whether he is still in pain, he will report that it is just as bad as ever, but that he really doesn't care. More evidence for dissociation of pain perception and pain reaction!

VI. Summary and Conclusions

From consideration of the nature of pain and a review of the literature, four main conclusions have emerged.

1. No operational definition of pain is at present available. The literature on pain has defined pain introspectively, and so far the definition has been quite satisfactory.
2. The psychophysics of pain has been well advanced as a result of sustained experimental attack in the last decade, but its further development depends on the correction of experimental errors and the use of more adequate statistical tools.
3. The localization of pain stimulation on the skin is a function of presentation to the central nervous system of a complex pattern of impulses from many different pain fibers of varying transmission speeds. Disturbances of this pattern will result in hyperalgesias and other disorders of pain perception. Qualitative differences in pain perceptions from different parts of the body are probably due only to differences in the pattern of impulses in the central nervous system resulting from differences in density of peripheral innervation.
4. Attention to the peripheral mechanisms of pain should not preclude recognition of the vital role played by the central nervous system in pain perception. In many cases, knowledge of what is happening at the periphery is inadequate for or unrelated to prediction of the pain experience. Study of cases of disorders of pain perception, especially causalgias, suggests that among the important processes going on in the spinal cord are (a) recruitment of fibers into an ongoing activity of internuncial neurons, either sustained by or independent of peripheral stimulation, (b) a vicious circle of pain-induced efferent activities resulting in sympathetic and somatic changes which accentuate the original pain stimulus, and (c) inhibition of pain by pain, pain by other skin senses, and possibly pain by the cortex.

Baltimore: Williams & Wilkins, 1939.

48. LEWIS, T. *Pain.* New York: Macmillan, 1942.

49. LIDDELL, H. S. Animal behavior studies bearing on the problem of pain. *Psychosom. Med.*, 1944, **6**, 261–265.

50. LIVINGSTON, W. K. *Pain mechanisms.* New York: Macmillan, 1943.

51. MCAULIFFE, G. W., GOODELL, HELEN, & WOLFF, H. G. Experimental studies on headache: pain from the nasal and paranasal structures. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 185–208.

52. MCLELLAN, A. M., & GOODELL, HELEN. Pain from the bladder, ureter, and kidney pelvis. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 252–262.

53. MICHELSEN, J. J. Subjective disturbances of the sense of pain from lesions of the cerebral cortex. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 86–99.

54. MOLONEY, J. C. The effort syndrome and low back pain. *J. nerv. ment. Dis.*, 1948, **108**, 10–24.

55. MORGAN, C. T. *Physiological psychology.* New York: McGraw-Hill, 1943.

56. PARSONS, C. M., & GOETZL, F. R. Effect of induced pain on pain threshold. *Proc. Soc. exp. Biol., N. Y.*, 1945, **60**, 327–329.

57. PARSONS, C. M., & GOETZL, F. R. The influence of acetylsalicylic acid upon analgesic effects of induced pain. *Permanente Found. med. Bull.*, 1946, **4**, 49–63.

58. PATTLE, R. E., & WEDDELL, G. Observations on electrical stimulation of pain fibers in an exposed human sensory nerve. *J. Neurophysiol.*, 1948, **11**, 93–98.

59. PENFIELD, W., & BOLDREY, E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*, 1937, **60**, 389–443.

60. RAY, B. S. The management of intractable pain by posterior rhizotomy. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 391–407.

61. REVICI, E., & RAVICH, R. A. The painful focus I. Physiological and pathological pain. *Bull. Inst. appl. Biol.*, 1949, **1**, 12–20.

62. REVICI, E., STOOPEN, E., FRENK, E., & RAVICH, R. A. The painful focus II. The relation of pain to local physico-chemical changes. *Bull. Inst. appl. Biol.*, 1949, **1**, 21–38.

63. RIDDOCH, G. Phantom limbs and body shape. *Brain*, 1941, **64**, 197–222.

64. ROBERTSON, S., GOODELL, HELEN, & WOLFF, H. G. Headache: the teeth as a source of headache and other pain. *Arch. Neurol. Psychiat., Chicago*, 1947, **57**, 277–291.

65. ROSENTHAL, S. S., & SONNENSCHEIN, R. R. Histamine as the possible chemical mediator for cutaneous pain. *Amer. J. Physiol.*, 1948, **155**, 186–190.

66. RUSINOV, V. Painful stimuli in causalgia. *Amer. Rev. Soviet Med.*, 1947, **4**, 436–438.

67. SCHILLING, R. F., & MUSSER, M. J. Pain reaction thresholds in psychoneurotic patients. *Amer. J. med. Sci.*, 1948, **215**, 195–197.

68. SCHUMACHER, G. A. The influence of inflammation on the pain threshold of the skin in man. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 166–172.

69. SHERMAN, I. C., & ARIEFF, A. J. Dissociation between pain and temperature in spinal cord lesions. *J. nerv. ment. Dis.*, 1948, **108**, 285–292.

70. SIMONS, D. T., DAY, E., GOODELL, HELEN, & WOLFF, H. G. Experimental studies on headache: muscles of the scalp and neck as sources

of pain. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 228–244.

71. SIMONS, D. T., & WOLFF, H. G. Studies on headache: mechanisms of chronic post-traumatic headache. *Psychosom. Med.*, 1946, **8**, 227–242.

72. SKILLERN, P. G. Clinical observations on: (I) Cutaneovisceral (somato-sympathetic) reflex arcs; (II) The role of hypermyotonia in bodily aches and pains. *J. nerv. ment. Dis.*, 1947, **105**, 449–464.

73. SLAUGHTER, D., & WRIGHT, F. T. A modification of the Hardy-Wolff-Goodell pain-threshold apparatus. *Anesth. & Analges.*, 1944, **23**, 115–119.

74. STONE, L. J., & JENKINS, W. L. Recent research in cutaneous sensitivity: I. Pain and temperature. *Psychol. Bull.*, 1940, **37**, 285–311.

75. STOOKEY, B. The management of intractable pain by chordotomy. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 416–433.

76. TAYLOR, J. Surgical treatment of pain. *Lancet*, 1938, **235**, 1151–1154.

77. THOMPSON, L., & GELLHORN, E. The influence of muscle pain on spinal reflexes. *Proc. Soc. exp. Biol., N. Y.*, 1945, **58**, 105–108.

78. TOWER, S. S. Pain: definition and properties of the unit for sensory reception. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 16–43.

79. TROTTER, W., & DAVIES, H. M. The peculiarities of sensibility found in cutaneous areas supplied by regenerative nerves. *J. Psychol. Neurol., Lpz.*, 1913, **20**, 102–150.

80. UGLOV, F. G. Novocaine block for injuries of the peripheral nerves. *Amer. Rev. Soviet Med.*, 1946, **4**, 68–69.

81. WALKER, A. E. Central representation of pain. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 63–85.

82. WALSHE, F. M. R. The anatomy and physiology of cutaneous sensibility. *Brain*, 1942, **65**, 48–114.

83. WATTS, J. W., & FREEMAN, W. Frontal lobotomy in the treatment of unbearable pain. *Res. Publ. Ass. nerv. ment. Dis.*, 1948, **27**, 715–722.

84. WEDDELL, G. The pattern of cutaneous innervation in relation to cutaneous sensibility. *J. Anat., Lond.*, 1941, **75**, 346–367.

85. WEDDELL, G. The multiple innervation of sensory spots in the skin. *J. Anat., Lond.*, 1941, **75**, 441–446.

86. WEDDELL, G., SINCLAIR, D. C., & FEINDEL, W. H. An anatomical basis for alterations in quality of pain sensibility. *J. Neurophysiol.*, 1948, **11**, 99–109.

87. WOLF, S., & HARDY, J. D. Studies on pain. Observations on pain due to local cooling and on factors involved in the "cold pressor" effect. *J. clin. Invest.*, 1941, **20**, 521–534.

88. WOLFF, H. G. Headache mechanisms—a summary. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 173–184.

89. WOLFF, H. G. Some observations on pain. *Harvey Lect.*, 1944, **39**.

90. WOLFF, H. G. *Headache and other head pain.* New York: Oxford Univ. Press, 1948.

91. WOLFF, H. G., & GOODELL, HELEN. The relation of attitude and suggestion to the perception of and reaction to pain. *Res. Publ. Ass. nerv. ment. Dis.*, 1943, **23**, 434–448.

92. WOLFF, H. G., & HARDY, J. D. On the nature of pain. *Physiol. Rev.*, 1947, **27**, 167–199.

93. WOLFF, H. G., HARDY, J. D., & GOODELL, HELEN. Studies on pain and the action of analgesics. *Trans. Ass. Amer. Physicians*, 1939, **54**, 325–328.

94. WOLFF, H. G., HARDY, J. D., & GOODELL, HELEN. Studies on pain: measurement of the effect of morphine, codeine, and other opiates on the pain threshold and an analysis of their relation to the

pain experience. *J. clin. Invest.*, 1940, **19**, 659–680.

95. Wolff, H. G., Hardy, J. D., & Goodell, Helen. Measurement of the effect on the pain threshold of acetylsalicylic acid, acetanalid, acetophenetidin, aminopyrine, ethyl alcohol, trichlorethylene, a barbiturate, quinine, ergotamine tartrate and caffeine: an analysis of their relation to the pain experience. *J. clin. Invest.*, 1941, **20**, 63–80.
96. Wolff, H. G., Hardy, J. D., & Goodell, Helen. Studies on pain: an analysis of the analgesic action of ethyl alcohol. *Trans. Ass. Amer. Physicians*, 1941, **56**, 317–320.
97. Wolff, H. G., Hardy, J. D., & Goodell, Helen. Studies on pain: measurement of the effect of ethyl alcohol on the pain threshold and on the "alarm" reaction. *J. Pharmacol.*, 1942, **75**, 38.
98. Wolff, H. G., & Wolf, S. *Pain.* Springfield, Ill.: Charles C Thomas, 1948.
99. Woolard, H. H., Weddell, G., & Harpman, J. A. Observations on the neurohistological basis of cutaneous pain. *J. Anat., Lond.*, 1940, **74**, 413–440.
100. Zikeyev, V. V. Causalgia and its treatment. *Amer. Rev. Soviet Med.*, 1946, **4**, 16–21.
101. Zotterman, Y. Touch, pain, and tickling: an electrophysiological investigation on cutaneous sensory nerves. *J. Physiol.*, 1939, **95**, 1–28.

Received May 13, 1950.

STATISTICAL TECHNIQUES IN PRESENT-DAY PSYCHODIAGNOSTICS[1]

IVAN NORMAN MENSH
Washington University Medical School

More and more the clinical psychologist is submitting the data he gathers to rigorous examination, applying the experimental principles of research design and statistical treatment to his everyday observations as well as to research material. This emphasis in present-day clinical psychology was well reflected in the symposium on Statistics for the Clinician (**29**) introduced by Zubin and presented to the Biometrics Section of the American Statistical Association. At that time a series of clinical problems was proposed by A. L. Baldwin, J. F. Kubis, L. S. Kogan and J. McV. Hunt, A. I. Rabin, and L. J. Cronbach. Solutions to the problems were suggested by W. Stephenson, E. F. Gardner, D. Horn, H. Guetzkow, and G. W. Brown. The present survey is much on the order of that symposium in reviewing the problems of statistical techniques in present-day psychodiagnostics and indicating the solutions attempted—whether proposed, planned, in progress, or completed.

The problems of validity, reliability, and other statistical treatment of data obtained in psychological practice have accompanied the use of diagnostic devices since their inception. As the projective methods became a more prominent tool of the clinical psychologist, still more statistical problems were posed because of the nature of the tests and the data obtained through their use. Within the past few months Cronbach (**6**) has reviewed the special area of the Rorschach method and presented evidence of the clinician's concern with "statistical confirmation for their hypotheses." The errors, omissions, and "unhappy choices of statistical procedures" were pointed out and suggestions made for more adequate statistical treatment. However, it is not only in the area of projective techniques that statistical questions have been raised; for these questions apply also to other psychodiagnostic devices. Loevinger (**14**) has pointed out some of the problems in scoring non-projective psychological tests of abilities, attitudes, and personality traits.

---

[1] A survey prepared for the Committee on Diagnostic Devices of the APA Division of Clinical and Abnormal Psychology, under the chairmanship of Dr. A. L. Benton, State University of Iowa. The author is grateful to Division members for their contributions (cf. Bibliography).

This report was presented in brief as a cross-committee study with the Research Committee of the Division, Dr. R. E. Harris, Chairman, at the 1949 American Psychological Association Convention, Denver, Colorado.

Cronbach recently outlined the area of diagnostic devices which is the concern of the present survey:

> In the past two decades, attention has been given principally to developing a workable tentative theory of personality and inventing promising diagnostic methods. Until one has materials worth validating, validation research must hold a secondary place. We now have projective techniques, situational tests, and numerous other new tools which purport to describe personality; there lies ahead a large chore of validation to define the limits of dependability of each technique (4, p. 373).

More specifically, the major emphasis in statistical treatment of data from the psychodiagnostic devices has shifted from "group-centered" to "individual-centered" statistics (27). This shift appears in the suggestions of Cattell (3), Horn (12), Stephenson (21, 22), Zubin (26, 27) and others. Nor is this entirely a new development, for a decade ago Woodworth (24) wrote of the need to combine statistical analysis with individual study, and of the advantages and disadvantages of the two approaches when applied separately. Even earlier, Allport (1) discussed "the study of general principles and the study of the individual case" and the philosopher Windelband's proposal to separate the *nomothetic* from the *idiographic* disciplines. Allport at that time pointed out that psychology "in the main has been striving to make of itself a completely nomothetic discipline," but cautioned that the two approaches should be regarded as overlapping and as contributing to one another.

Zubin (27) felt that the recent symposium on Statistics for the Clinician was, in fact, "an attempt at quantifying the idiographic approach," but recognized that in present-day practice the entire emphasis is not on the individual case, because in that symposium "not all the enthusiasm of the present writer for 'individual' statistics is shared by the other contributors." Yet the stress on the individual case and on "intra-individual" study is very much evident in current literature. Why the trend toward individual study, particularly by the clinician?

Basically and ultimately, of course, the clinician is concerned with the individual case, for the single person is the unit in psychodiagnostics, in therapy, and in research, no matter how many units are studied. But the problem goes beyond this, and with the exception of relatively recently developed statistical techniques there have been no other than the methods based on the assumptions of normal distribution of a series of cases. This does not mean, however, that the "normal law of error" of Laplace, Gauss, Quetelet, and Galton has gone unchallenged. Particularly since the introduction of the projective methods have questions been raised about the normal distribution curve and its application to data obtained by means of such psychodiagnostic devices. In the early days of the use of the Rorschach method in this country, Meltzer (**16**) noted that "the habit of regarding the normal distribution curve as a point of reference" meant that the psychologist saw test results which deviated from the normal curve as "indicative of a poorly constructed test." Meltzer emphasized that it was not necessary to assume that all measures of human nature must yield normal distributions, and cited the discussions in E. L. Thorndike's *Mental and Social Measurements* and T. L. Kelley's *Statistical Methods* in evidence.

Recently this argument again has been advanced by a number of investigators. Rabin (**18**) summarized, as one of the statistical problems faced in Rorschach analysis, that there are "*n* factors which are not similarly distributed; some follow a normal distribution, some a *J*-curve and some other forms not yet investigated." Gardner (**9**) also has written that there are "many situations in which the assumption of normality seems unwarranted . . . it often seems more reasonable to expect a skewed distribution rather than a normal distribution. . . . "

Finally, many students of the problems of statistical methods for data from psychodiagnostic devices are now proposing re-evaluation of our thinking in this area. These proposals principally have concerned the projective techniques, although the more traditional devices also have been re-examined in this light. Loevinger has summarized the trend:

> In the field of measuring personality characteristics, there has been far from unanimous agreement as to the value of traditional psychometrics, and some of the most promising tests, notably the projective tests, have largely bypassed traditional procedures. Even with reference to measuring abilities, the pages of *Psychometrika* have contained an increasing number of articles in the past few years exploring the contradictions and dilemmas which arise from application of the concept of reliability, of the method of rectilinear regression, and of heterogeneity of test content. . . . Development of factor analysis has sharpened the appreciation of these difficulties (**14**, p. 521).

Further, Symonds (**23**) prefaced his recent text on personality dynamics with a statement that the demand by the scientific method for the reproducibility of data cannot now be met in clinical practice. He emphasized the difficulties facing an attempt to "repeat all of the conditions which would be necessary" in order to obtain the same data in successive samples in psychotherapy for reliability and validity studies. Also, there are no absolute definitions of "normality and good adjustment," for the criteria vary according to the culture, the place, the time, and the individual—"Each person makes the only adjustments possible for him. . . . " (23, p. 569). This is the philosophy of the ultimate in the idiographic approach. Frank's position has been similar in nature, and Cronbach (7) concludes that he "appears to argue that cases cannot be treated objectively, quantitatively, and anonymously, when projective tests are used." Yet, there are those, like Grant (10), who feel that the basic concepts are as appropriate for clinical data as for other material and that the problem is one of devising special tests and techniques for treating the obtained data. In reviewing "new procedures of experimental design and analysis," Grant finds many which already have proved useful for clinical problems. Moreover, it is his feeling that "The theory of experimental design has, at least temporarily, far outstripped psychological practice."

PROPOSALS FOR SOLVING THE STATISTICAL PROBLEMS IN PRESENT-DAY PSYCHODIAGNOSTICS

Zubin (27) already has summarized a number of the methods for treating data from diagnostic devices, including pattern analysis, factor analysis, other correlation techniques including both the traditional *R*- and the newer *Q*-technique or inverted-factor analysis, analysis of variance and covariance, rating and scaling, and the Fisher discriminant function. In Grant's extensive review (10) he covers experimental design and analysis, psychometric methods, correlation and multivariate analysis, coding and computational techniques, and bibliographies and books for general treatments. The methods include analysis of variance with both inter- and intra-subject variation accounted for, the determination of sample size in experimental design, the interpretation of the extreme deviate or the Studentized u, the effects of variation in criteria on the *F*-test, Yates' contingency methods, non-parametric techniques when the assumptions of normality appear to be unwarranted, scaling methods improving on Guttman's analysis, Grubbs' method for distinguishing measuring instrument inaccuracies from actual product variability, Horst's generalized expression for the reliability coefficient when the number of measures varies from individual to individual, and multivariate tests. To these may be added the *P*-technique to round out the *P-Q-R* triumvirate, interval scaling, homogenous test construction, validation by successive clinical predictions, and matching.

Pattern analysis. Munroe has emphasized that statistical treatment of test data must capitalize on clinical experience and specifically has urged consideration of combinations of scores in quantifying observations.

> By and large it seems clear that poor results [of quantitative studies of Rorschach data] are obtained when a limited number of the traditional scores (which were never intended for use in isolation) are equated with personality traits without further ado and correlated with external factors. Promising results are obtained when *combinations* of scores are used . . . (**17**, p. 15).

Cronbach (**7**) also has pointed to the inadequacy of conventional treatment of test scores. He acknowledges the simplicity of handling test data by treating one score at a time, but is concerned with the failure of this technique to take into account the interrelations and interactions of scores. Yet there is the problem of describing the many dimensions of personality and considering all the data simultaneously. Cronbach suggests, "If tests could be made more reliable, equally reliable, and equally intercorrelated, clinical testing would more quickly produce significant findings." Here, then, are at least three goals to be pursued.

Recognizing the practical limitations of working with a great number of dimensions of personality, Cronbach (**5**) offers the alternative of reducing this number by the "signs approach" of many investigators, or by the relatively newer multiple regression and discriminant function methods. After discussing the limitations of these statistical treatments in terms of the amount of information lost in the course of the analysis and the involved functions, he proposes pattern tabulation. Even this method is limited in its narrow scope of dealing with only two or three scores at a time. Yet it does permit the examination of three-dimensional data without the need for elaborate statistical procedures and as such is an advance over the handling of but one score at a time—"The technique permits one to summarize the performance of a group of persons on a pattern of several scores, to compare groups and to test differences in pattern for significance."

In a sense, this technique is a reversal of the usual trend of applying group statistics to intra-individual data, for pattern analysis of the latter is well-founded in the study of the individual and now Cronbach proposes an application of the method to group data.

Rabin (**18**), in tracing the history of the Rorschach method in differential diagnosis, also noted the inadequacies of the "signs approach" and concluded that it gave "hints," but not differential diagnosis. However, Zubin (**26**) used pattern analysis over ten years ago and very very recently reaffirmed his belief in its promise as a method useful for psychodiagnostics.

Sells[2] has been developing at the USAF School of Aviation Medicine an approach which "utilizes scoring symbols and devices which have been found useful clinically and . . . fit into simple schemes for quantitative treatment." Applied to psychiatric screening tests, the technique classifies behavior as satisfactory, questionable, or unsatisfactory. For example, in treating group Rorschach data and scoring for neurotic tendency, satisfactory behavior consists of no rejections of cards, 20 or more *R*, 30–40% *A*, no color naming, etc.; questionable behavior includes one rejection, 10–20 *R*, *A*% greater or less than 30–40, one instance of color naming, etc.; and unsatisfactory records are defined as more than one rejection, fewer than 10 *R*, greater than 50 % *A*, more than one instance of color naming, etc. Such scoring is based upon discriminating power of the various indicators (the "signs approach" is seen in Sells' technique) and "empirical examination." The various items may be given relative weights and placed in combinations (patterning) on the basis of (1) discriminating power and, (2) relative importance as defined by frequency of occurrence of the item.

---

[2] Sells, S. B. *Quantitative treatment of projective test data.* 1949 (unpublished).

Factorial analysis. Factorial analysis has been widely used in the treatment of various types of test data. Rosenzweig summarized its role as follows: "While not a method of directly appraising or investigating a personality. . . . By factor analysis quantitative data are treated mathematically to uncover the smallest number of basic variables to which the crude observations can be reduced, and to throw light upon the relationships among these variables" (**19**, p. 359).

Eysenck (**8**), using factor analysis in his studies of the dimensions of personality, points out that the method, historically, is an extension of the "method of concomitant variation" underlying all correlation procedures. Four different types of factors can be developed—general, group, specific, and error factors. However, there is little agreement on the meaning of these factors after they have been isolated, and Eysenck lists interpretations which vary according to the investigator:

> . . . Thurstone (1935) and Holzinger (1937), for instance, regard factors as primary or fundamental abilities; similarly, Spearman (1927) regards them as fundamental functions of the mind. If factor analysis is to be applied to the study of personality as well as to the study of abilities, this definition has to be broadened somewhat, and it would become necessary to think of factors as elementary or unitary traits of personality (Kelley, 1935), or as the fundamental dimensions of the mind (Guilford, 1940). On the other hand, Anastasi (1938), Allport (1937), and Thomson (1939) consider that factors are statistical artifacts, having no "reality" of any kind.
>
> Our own position is very similar to that of Burt (1940), who regards factors as principles of classification (**8**, pp. 16–17).

Thurstone, Guilford, R. B. Cattell, and Eysenck, among others, have used factor analysis to extend the amount of information derived from a set of data. We see the traditional correlational techniques built upon to form such analyses as Cattell's "*P*-technique" and Stephenson's

"*Q*-technique" (**3**, **21**, **22**). However, the need for more precise definition and for more universal agreement on the *meaning* of factor analysis is again underlined in Stephenson's presentation of his *Q*-technique, just as in Eysenck's comments. Stephenson (**21**) cites Banks' use of Burt's summation method to refactor a table of correlations which previously had been reduced to eighteen factors by Cattell, and found a different set of factors involving one general, two group, and eight other factors. Stephenson, in turn, refactored the Cattell table by Spearman methods and found four group factors. Further, Stephenson summarizes: (1) Banks considered her factors were structurally related, (2) Cattell's were not, and (3) Stephenson's interpretation of the factors were "very different."

Nevertheless, Cattell has made factor analysis an integral part of his treatment of data in personality study and through such analysis recently has built much of his personality theory. For example, in his exposition of the *P*-technique, the method of factor analysis is the basic structure on which, together with the coefficients which are factored, the technique rests. But the weight of evidence still argues for more universal understanding among those who use factor analysis, if it is to satisfy many of the statistical problems raised by present-day psychodiagnostics. This criticism is emphasized by Loevinger (**14**) who noted that "A test which is factorially pure in one system [of factor analysis] will not be so according to others." There is very much the need for a common language in factor analysis as in other techniques for treating test data.

P-Q-R-techniques triumvirate. Although the traditional correlational technique, the *R*-technique, is more than a half-century old, only very recently have there been adaptations designed for the specialized problems arising from personality study as we know it today. Thus, there is as recently as 1947, Cattell's article (**3**) which in its summary begins "*P*-technique, a method employing intra-individual correlation, is tried out for the first time." And Stephenson (**21**) dates to a decade ago his first application of the *Q*-technique. Cattell in a footnote distinguishes the *P*-, *Q*-, and *R*-techniques thus:

> . . . in *R*-technique we correlated *test variables* with regard to a series of persons; in *Q*-technique we correlate *persons* with regard to a series of tests. In *P*-technique we again correlate test variables, but with regard to a series of occasions and within a single person. Practically all factor analyses yet published have been in terms of *R*-technique and the remainder in *Q*-technique (**3**, p. 267).

The method of "temporal sequence studies," implicit in the *P*-technique and using intra-individual correlation, has been mentioned by a number of investigators other than Cattell. Nor is the method restricted to correlational techniques, as pointed out by Horn (**12**) who proposed three methods for treating intra-individual variability in the extending the method of scaling to a variety of psychodiagnostic devices. Thus, they have applied the "psychometric approach" in the evaluation of such techniques as the Rorschach, Levy Movement Blots, the TAT, and word association. In preparation there is a volume to include scaling of handwriting analysis, the mosaic test, and the Bender-Gestalt. An example of the scaling technique for "Perseveration Tendency (Pe)" follows:

0. Response totally independent of previous responses.
1. Some slight or general similarity either in form or content to a previous response.
2. Definite similarity in form but not in content.
3. Definite similarity in content but not in form.
4. Exact, or nearly exact, repetition of a previous response.

Such scales, of four or five steps, are defined for Rorschach responses with respect to locale, organization, determinants, interpretation (by the subject of the determinants used by him), content, psychological characteristics of the response (as perseveration, mood, congruence), subject's attitude toward the response, and general data (including number of rejections, inquiry responses, and reaction time). These in turn consist of a series of scales. To define locale, for example, there are scales of importance, type of white space, inclusiveness of primary locale, and frequency of locale. It is obvious from the scope of the scaling by Zubin and Young that such quantification of projective data represents a major effort toward the development of this type of "raw statistical tool."

In a different area of research—social service case work—but with the same emphasis on quantification and reliability of judgment is the work of Kogan and Hunt. These investigators (**13**) also used the rating scale in the primary quantification of data, constructing one by psychological principles and including anchoring case illustrations. Objectivity was thus furthered by having each judge rate by applying the standardized scale to a structured case summary. But Kogan and Hunt report, "In the course of the development and application of this scale a persistent statistical problem has confronted the investigators." Working through the problems in terms of null hypotheses and the broader problem of evaluation, they conclude that their "critical problem is how best to determine whether alternative methods designed to increase agreement among judges really do have such an effect." After using judgments of raters as the raw data in correlations and in *t*- and *F*-tests of significance, final evaluation showed their dissatisfaction with any one of these methods for treating multi-judge reliability.

Here again, the problem seems to lie in the further quantification and statistical treatment of data, with use of the rating scale remaining a satisfactory initial tool for quantification. This further quantification and treatment recently has been developed by Yates (**10**), who now has methods for treating contingency table problems where the psychologist has data only roughly quantified by rating techniques.

Coding. Similar to the rating scale in that the technique makes data available for statistical treatment is the method of coding. Guetzkow (**11**) recently has attempted to make explicit some of the thinking on coding. Not only does he present the procedures for transforming qualitative material into unitized and classified data, but he examines the reliability of the coding procedures. After indicating the two operations in coding—separating the qualitative data into units, and establishing "category-sets" for classifying the units—Guetzkow details the frames of reference, dimensionalities, and interrelationships of category-sets. He is careful to recognize that unitizing of data is a problem in coding, in order to avoid the common error of failing to delineate unitizing from the data collection procedure and the subsequent classification process. Different unit levels may be used simultaneously, just as there may be simultaneous use of several category-sets at different levels of generality in order to preserve the richness of the data. It can be seen here how coding resembles the rating scale as a statistical tool for quantifying data, for it will be recalled that Zubin urged scaling as a means "to capture the full clinical flavor of the response" (cf. p. 483, above).

Reliability in coding, Guetzkow specifies, depends upon the accuracy with which the data have been coded, and upon the sample (whether its size and selection are adequate to characterize reliably the behavior studied). The latter is easily recognized as a problem in sampling, and not specific to coding, whereas accuracy of coding is as significant as accuracy in obtaining the raw data. Reliability estimates are developed by Guetzkow in a series of equations for both categorizing and unitizing. Reliability of categorizing involves principally the following: "The proportion of units upon which two coders [the reliability of more than two coders also is evaluated by Guetzkow, so that restriction in number is not necessary] agree may be conceived as the sum of those items which both coders incorrectly classify in the same incorrect way."

The equations for evaluating the reliability of the unitizing of data are based upon comparison of the number of units obtained by the coders. Of further practical importance is the contribution of these reliability estimates to the investigator's decision as to how large an amount of data needs to be check-coded in order to obtain the desired level of accuracy. This contribution calls to mind Zubin's suggestion (**27**) that "sequential analysis provides the experimenter with a means of determining how many observations are required for testing a given hypothesis."

Interval scales. In quantifying data, rating scales basically define the characteristics of the dimensions scaled. This function of definition has often been the concern of the psychologist. As Gardner (**9**), Loevinger (**14**), McNemar (**15**), and others have pointed out, scaling methods have been widely used in studies of attitudes, abilities, aptitudes, and other personality factors. Gardner (**9**) distinguishes between ordinal scales (so called by Stevens) and interval scales (the units of which are equal throughout the range of the scale), quoting McNemar's argument that ordinal scales give only relative rank order among individuals. Gardner writes that interval scales furnish two kinds of information not otherwise available.

> Although ordinal scales are adequate when we wish to know whether one person is superior to another or simply whether improvement—regardless of amount—has taken place, they are inadequate in two important respects. In the first place they do not permit the determination of the *amount* of growth of an individual in a particular trait. Nor do they, in the second place, permit the comparison of differences in performance of individuals in a particular trait. To provide these two kinds of information, interval scales . . . are needed (**9**, p. 39).

Further, Gardner proposes a method for developing an ordinal scale in which the assumption of normal distribution is not necessary. It will be recalled (cf. pp. 477 ff.) that this assumption is one which the clinical psychologist often feels is not warranted. Gardner also is of the opinion that it is often more reasonable to expect a skewed distribution than a normal distribution. In transforming an ordinal scale of raw scores from two school grades into an interval scale he fits to the data "the more general Pearson Type III Curve," for which he gives the equation. The method has been applied to achievement scores on several parts of the Stanford Achievement Tests taken by approximately fifty thousand children in Grades 2–9. Gardner sees his interval scale method as useful in the measurement of growth and in the comparison of differences where a skewed distribution seems more likely than the usual assumption of normal distribution.

Homogeneous tests. Setting up as her goal the measurement of one psychological function at a time, Loevinger (**14**) has developed the technique of "homogeneous tests."[3] She distinguishes between *cumulative homogeneous tests* and *differential homogeneous tests*. In the former, "when the items are ordered according to decreasing number of pluses, each person scores plus up to a characteristic item and minus on all subsequent items. For *differential homogeneous tests*, there is an order of the items such that each individual scores minus up to a characteristic item, plus up to another characteristic item, and minus on subsequent items" (**14**, p. 527). Loevinger gives in example a measure of emotional maturity. "Laces own shoe," "goes downtown alone," "uses family car" are suggested by her as items of the cumulative test variety; and preferred types of books, radio programs, and games would represent differential test items.

---

[3] It is interesting to note that this same goal has been explored very recently by another method. The Levy Movement Blots (**28**) have been designed "to reduce the components of the Rorschach stimulus."

The motivation for developing the technique of homogeneous tests seems best represented in Loevinger's criticism of psychological tests of abilities, attitudes, and personality traits. She summarized her conception of the present state of affairs as follows:

> In contemporary psychometric practice, it is the rule rather than the exception that two people having the same score on a test will have scored plus on different items. When each score stands for a variety of patterns of pluses, it is difficult to conceive of these scores as measuring anything. Such scores are crude empirical devices known to have some predictive efficiency, but they cannot be called measurements in any strict sense (**14**, p. 507).

Successive clinical predictions. Rosenzweig (**20**) has adapted from the successive approximations approach a method which he terms "successive clinical predictions." Emphasizing the lack of quantitative definition of external criteria of validity of psychodiagnostic devices, he reviews the usual criteria of validity—qualitative and quantitative judgment, correlation with other diagnostic devices, and matching data from two or more sources—and points out the problems accompanying the use of such criteria, including rating errors, lack of "systematic validation of referrent instruments," and the presumptive validity of matching.

The framework of Rosenzweig's proposal is familiar to the research worker. His "method for validating projective techniques projectively" he defines as, "a systematic use of trial and error with a succession of hypotheses being advanced, checked, rejected or revised until finally the character of the instrument has been defined by empirical investigation." The method is implemented by a seven-step procedure combining the judgments of two or more investigators; the administration of the new test and a variety of referrent ones; evaluations from case history, interview, test, and other material sources; predictive attempts, qualitative and quantitative; degree of agreement between predictions and test data; revisions of the hypotheses according to the information gathered in earlier steps; and repetition of the steps with the revised hypotheses. Rosenzweig concludes: "A succession of such experiments in clinical prediction should make it possible for the new device gradually to define its own validity—a procedure of objective self-definition."

Method of the discriminant function. Zubin (**27**) sees promise in the method of the discriminant function as a statistical tool. He defines the technique as one "which serves to integrate a group of variables into a total score such that the maximum differentiation from a contrasted group is effected." Moreover, a very recent development of partial discriminant functions, related to zero order discriminant function as partial correlations are to zero order correlations, may permit the comparison of two groups postoperatively on a test on which they had not been measured and equated preoperatively.

Examining a problem the converse of the Fisher discriminant function, Zubin (**27**) also reports a need of a method for discovering psychological types by "fractionating a given population into such sub-groups that the difference between the means on these sub-groups would be a maximum." Stephenson feels that his approach, the *Q*-technique, may be used for determining psychological types by correlating each subject with all the others and factor analyzing the resultant matrix (cf. p. 483, above).

Matching. Since Vernon's application of the blind-matching technique to personality research in the 1930's, this method of treating qualitative data has been widely used (**4**, p. 367), particularly for interpretations of projective test data. "Blind analyses" are blind in the sense that the personality interpretation is made without actually seeing the patient whose personality is under study.

Just as Zubin (**26**, **28**) has reaffirmed his faith in rating scales for handling qualitative data, so has Cronbach (**4**) experimented again with the blind-matching technique. Rosenzweig (**19**, p. 357 ff.) has attributed the popularity of this method as partially a function of the emphasis on the "total individual." He sees as the basic motivation for the development of the matching technique "a comparison of individuals and at the same time . . . regard for the interdependence of the personality in its various aspects . . . as opposed to the rating scale method—no attempt is ordinarily made to break down the personality into segments or parts regarded as more or less independently measurable."

Cronbach (**4**) objectively reviews the arguments *pro* and *con* blind-matching, weighs the evidence, and concludes from the application of the method to a small study that "It yielded the types of objective analysis necessary for a thorough validation of a qualitative assessment of personality." The procedure for handling three cases at a time in matching is outlined in five steps and then detailed by Cronbach. He concisely presents the advantages and disadvantages of the method.

> The advantages of the method may be summarized as follows:
>
> 1. Neither the bias of the investigator nor of the judges influences the estimate of validity.
> 2. The estimate of validity, although using a qualitative criterion, yields a test of statistical significance and a coefficient like conventional validity coefficients.
> 3. Qualitative criteria are more meaningful and more truthful than ratings or other quantitative criteria.
> 4. The method identifies objectively the accurate and inaccurate aspects of the prediction.
> 5. The method permits one to identify the types of cases for whom prediction is relatively accurate or inaccurate.
> 6. The disadvantage of the method is that the design is laborious, and the

services of a large number of judges must be enlisted. Since such a large group of judges cannot be handpicked, there is some danger of careless judging which will reduce validity. The difficulty of obtaining adequate qualitative criteria must be recognized, but this problem is not specific to this validation design. The present method is not ideal, since it does not treat the prediction as a whole; but blind-matching, which does treat the prediction as a whole, gives little or no information about the validity of the parts (**4**, p. 373).

Finally, Grant (**10**) has reviewed a special technique which he sees as promising in the use of the matching method. He describes Levene's derivation of a sampling distribution of matches which may be applied in the test of the diagnostic qualities of a complex test, by matching the protocols or scores with the diagnostic group for a number of patients where the judge is told the size of each diagnostic group.

EXPERIMENTAL DESIGN

The previous section on statistical techniques has repeatedly introduced experimental design as a major factor in the development and/or application of techniques. Statistics and design are complements in research, as Grant (**10**) has shown in his review on "Statistical Theory and Research Design." The emphasis in the past pages has been on the statistical techniques, however, and Grant's emphasis is on the second of the two components—research design.

Of interest to the clinician, in addition to the "Proposals" just presented, are certain procedures selected from Grant's survey. For example, the work of Harris, Horwitz, and Mood provides methods that are useful in the problem of size of sample in research. In another area, Nair developed a technique for evaluating an extreme deviate mean, the Studentized u, one found in analysis of variance to differ extremely from a group of means. Grant points out the difficulty frequently found in the interpretation of such a deviate, and sees in Nair's treatment "an objective test for rejection of outlying observations or the evaluation of an unusual mean." Also valuable in analysis of variance is Aspin's demonstration that in a two-sample experiment, absence of homogeneity of variance need not be serious if the two samples are of equal size.

Although there is additional freedom provided by the above techniques, there also is the caution by Grant—of particular note to the clinician who proposes such research—that the plan must include precise control of the independent variable in analysis of variance designs. Madow's example of the sensitivity of the F-test to lack of control of the independent variable is elaborated in Grant's suggestion: " . . . if three differentially diagnosed groups of patients were selected for evaluation of improvement after four different therapies, slight variation in diagnostic criteria or in details of therapy could greatly reduce the power of an F-test used in analysis" (**10**). Also because of the clinician's attention to diagnosis and prediction, where the samples he studies may differ in two or more variables the problem of multivariate analysis must be faced. Analysis of variance (**10**) provides tests of significance of differences between groups when one variable is involved, but does not yield tests for differences "in two or more variables, and a test is sought for detecting systematic differences between groups, making use of all available scores." Grant's review of Rao's, and of Rao and Slater's multivariate analysis indicates the value of the method for inter-group analyses.

SUMMARY

Psychodiagnostic devices in clinical practice have contributed significantly to the study of human behavior. Concerned with techniques for establishing the reliability and validity of the diagnostic tools, psychologists recently have faced problems in treating data which have shifted the thinking of the statistically-minded. The projective techniques especially have presented problems.

Traditionally, the assumption of normal distribution has furnished the basic hypothesis for statistical treatment of test data. Empirical evidence has been evaluated by a number of investigators (**2**, **13**, **14**, **17**) who have concluded that the assumption of normal distribution often is unwarranted. Although the history of statistics in psychology has emphasized the treatment of group data, the emphasis has now shifted to techniques designed to treat data from a single individual in an "intra-individual" (**2**) or "individual-centered" (**27**) approach.

Another direction of thinking, and more basic, is one which concerns the *meaning* of statistics for test data, particularly from projective methods. There are those who feel that the underlying assumptions of statistics in psychometrics do not apply to projective test data. Other investigators, however, feel that the assumptions apply equally well to all test data, and that the major problem is the devising of special techniques to fit the data of projective devices.

There is this lack of agreement on the problems of handling test data statistically, yet there has been progress. This progress is seen in a number of statistical techniques proposed—some dealing with the quantification of data so that they can be ordered for statistical treatment, e.g., rating scales and coding; others concern methods for treating data already ordered in codes, scales, correlations, or other distributions. These include the more familiar correlational or *R*-technique, and factor analysis of the correlational data. Techniques proposed more recently include the *P*- and *Q*-techniques, pattern analyses, discriminant and partial discriminant functions, interval scaling, homogeneous test construction, successive clinical predictions, matching, analysis of variance and multivariate analysis.

Despite the lack of agreement among the various investigators, it is significant in the current emphases of the clinical psychologist that he is more and more concerned with rigorous examination of his techniques and of his methods for treating the data furnished by the diagnostic devices. Within the past two years psychological meetings have featured symposia titled "The problems of quantification and objectification in personality" (**25**) and "Statistics for the clinician" (**29**), as witness of the importance of these problems to the student of personality. Such interest and effort are encouraging research toward the goals of reliability and validity in the study of personality.

BIBLIOGRAPHY

1. ALLPORT, G. W. *Personality: A psychological interpretation.* New York: Holt, 1937.
2. BALDWIN, A. L. Statistical problems in the treatment of case histories. *J. clin. Psychol.*, 1950, **6**, 6–12.
3. CATTELL, R. B., CATTELL, A. K. S., & RHYMER, R. M. P-technique demonstrated in determining psychophysiological source traits in a normal individual. *Psychometrika*, 1947, **12**, 267–288.
4. CRONBACH, L. J. A validation design for qualitative studies of personality. *J. consult. Psychol.*, 1948, **12**, 365–374.
5. CRONBACH, L. J. "Pattern Tabulation": A statistical method for analysis of limited patterns of scores, with particular reference to the Rorschach Test. *Educ. psychol. Measmt*, 1949, **9**, 149–172.
6. CRONBACH, L. J. Statistical methods applied to Rorschach scores: A review. *Psychol. Bull.*, 1949, **46**, 393–429.
7. CRONBACH, L. J. Statistical methods for multi-score tests. *J. clin. Psychol.*, 1950, **6**, 21–26.
8. EYSENCK, H. J. *Dimensions of personality.* London: Kegan Paul, Trench, Trubner, 1947.
9. GARDNER, E. F. Comments on selected scaling techniques with a description of a new type of scale. *J. clin. Psychol.*, 1950, **6**, 38–43.
10. GRANT, D. Statistical theory and research design. In C. P. Stone (Ed.), *Annual review of psychology*, 1950, 277–296.
11. GUETZKOW, H. Unitizing and categorizing problems in coding qualitative data. *J. clin. Psychol.*, 1950, **6**, 47–58.
12. HORN, D. Intra-individual variability in the study of personality. *J. clin. Psychol.*, 1950, **6**, 43–47.
13. KOGAN, L. S., & HUNT, J. McV. Problems of multi-judge reliability. *J. clin. Psychol.*, 1950, **6**, 16–19.
14. LOEVINGER, JANE. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol. Bull.*, 1948, **45**, 507–529.
15. McNEMAR, Q. Opinion-attitude methodology. *Psychol. Bull.*, 1946, **43**, 289–374.

16. MELTZER, H. Personality differences among stutterers as indicated by the Rorschach Test. *Amer. J. Orthopsychiat.*, 1934, **4**, 262–282.

17. MUNROE, RUTH L. Projective methods in group testing. *J. consult. Psychol.*, 1948, **12**, 8–15.

18. RABIN, A. I. Statistical problems involved in Rorschach patterning. *J. clin. Psychol.*, 1950, **6**, 19–21.

19. ROSENZWEIG, S. Available methods for studying personality. *J. Psychol.*, 1948, **28**, 345–368.

20. ROSENZWEIG, S. A method of validation by successive clinical predictions. *J. abnorm. soc. Psychol.*, 1950, **45**, 507–509.

21. STEPHENSON, W. The Q-technique study of personality. *Trans. N. Y. Acad. Sci.*, 1949, **11**, 215–219.

22. STEPHENSON, W. A statistical approach to typology: The study of trait-universes. *J. clin. Psychol.*, 1950, **6**, 26–38.

23. SYMONDS, P. M. *The dynamics of human adjustment.* New York: Appleton-Century, 1946.

24. WOODWORTH, R. S. Fundamental and applied research. In *Symposium on mental health.* Lancaster, Pa.: Science Press, 1939.

25. ZUBIN, J. Introduction. The problems of quantification and objectification in personality: A symposium. *J. Personality*, 1948, **17**, 141–145.

26. ZUBIN, J. Personality research and psychopathology as related to clinical practice. Clinical practice and personality theory: A symposium. *J. abnorm. soc. Psychol.*, 1949, **44**, 14–21.

27. ZUBIN, J. Introduction: Symposium on statistics for the clinician. *J. clin. Psychol.*, 1950, **6**, 1–6.

28. ZUBIN, J. & YOUNG, K. M. *Manual of projective and cognate techniques.* Madison, Wis.: College Typing, 1948.

29. Symposium: Statistics for the Clinician. *J. clin. Psychol.*, 1950, **6**, 1–76.

Received January 13, 1950.

RORSCHACH VALIDATION: SOME METHODOLOGICAL ASPECTS

LEONARD I. SCHNEIDER
The State College of Washington

Introduction

It is generally agreed in clinical circles that the Rorschach psychodiagnostic method is the most comprehensive device in the clinician's repertoire for ascertaining personality patterns. However, it is also agreed that results are highly dependent upon the skill of the interpreter. This is due to the fact that the relationships between the data yielded by the test and personality variables have not been clearly ascertained and stated. Consequently, the user of the method must rely upon a body of guesses as to the relationships involved—including those which have accumulated in the literature, plus his own experience with the test in particular, and knowledge of normal and abnormal personality in general.

Considerable space in psychological literature was at one time devoted to the question of whether or not such a state of affairs was desirable. More recently the majority of those concerned with use of the instrument have supported the position that the Rorschach needs validation (**1**, **2**, **6**, **8**, **15**, **16**, **17**, **21**, **26**, **30**, **32**, **35**, **36**, **39**). However, the variety of opinions on the requirements for Rorschach validation are as numerous as are the theoretical and methodological positions in the entire field of psychology—as the following quotations will demonstrate.

Harrison (**15**) claims that rigid validation procedures may be applicable to cognitive aspects of the personality but not the affective aspects. He maintains that "there must be sacrifice of objectivity and spurious precision with more reliance being put upon insight and ingenuity in the forging of tools for personality investigation." Clinical validation, he claims, is ample evidence for the present, though controlled experimentation may in the future produce more definite knowledge. Rapaport (**26**) and Brosin and Fromm (**6**) call for experimental investigation of "what form level means" by use of Gestalt experiments of form genesis and prägnance. Brosin and Fromm state further that ". . . the question of what may constitute the basis for correspondence between the responses and the personality of the subject; and the nature of the equivalence between the perceptual patterns and the general personality patterns, seem to obtain the most fruitful suggestions from Gestalt Psychology."

Sargent (**32**) claims that "factors taken out of context have little meaning. For the reason that it is not the absolute amount of one determinant but its relation to the whole pattern which gives it significance in the individual protocol." Thus, for her, validity studies can proceed along two lines: (1) correspondence of Rorschach impression of the total personality with impressions independently arrived at, and (2) predictive capacity which she thinks is the sounder approach due to the fact that correlations based upon inter-judge agreement are minimized by "each analyst understanding correctly different aspects of the personalities under inspection."

Macfarlane (**21**) lists five validation methods: (1) correlation with outside criteria, (2) projective data vs. case history data, (3) through time consistencies, (4) collateral experimental approach, and (5) degree of success in prediction, which she offers as the most fruitful method. In support of this method she says:

> The writer's opinion is that the utilization of the interpretation and predictive judgments of widely experienced clinicians later checked for predictive success, will offer at this stage the most productive leads. If an experienced clinician is able to predict with considerable success, then the data on which he bases his correct intuitive predictions can be inspected and validly weighted configurations can be established, quantified, and made available to less experienced people. Also, his wrong and partially right predictions in conjunction with the right ones can serve for finer and more differentiating criteria.

Frank (**8**) holds that validation must be concerned with the aggregate of part functions since he believes that there may be lawfulness at that level but not between the discrete part functions and outside criteria. Like Harrison, he appears to believe in the acceptability of intermediate gross validation, for he says:

> If it appears that the subject projects similar patterns or configurations upon widely different materials and reveals in his life history the sequence of experiences that make these projections psychologically meaningful for his personality, then the procedures may be judged sufficiently valid to warrant further experimentation and refinement.

Thurstone (**35**), in a recent polemic against the scientific laxity of Rorschach workers, calls for validation data that will enable a Rorschach interpreter to predict the style of response an individual is likely to give to different types of life situations. Hertz (**16, 17**) has been quite outspoken about the need for validating data. She suggests four methods: (1) direct experimentation, (2) comparison with extensive individual case studies, (3) comparison with independent, objective data, and (4) comparison with known diagnoses and clinical pictures. In her earlier article (16), she states that concentration on one individual in the tradition of the clinician would be scientifically productive.

Varvel (36) has offered some specific suggestions for validation procedures. He mentions the possibility of finding the relationship between the autokinetic phenomenon and Rorschach personality types, checking Rorschach factors against data obtained from the Dembo situation, use of drugs and post-hypnotic suggestion to measure the effects on Rorschach protocols of these experimentally produced states, and simultaneous recording with the Luria technique, pneumograph and Rorschach responses.

In 1935 Beck (1) wrote that the value of the Rorschach method was not on the basis of successful penetration into personality as yet, but as a suggestion that human nature could be passed through a prism and analyzed into component elements. He cautioned, however, that it must be recognized, " . . . that the same trait is not necessarily equivalent in two different kinds of personalities, or in another way, the same diagnostic index may have different diagnostic value depending on the larger background of the personality in which it appears." In 1942 (2) he wrote that Klopfer's argument that standardization of the Rorschach implied devising of numerical formulae for the several whole personality entities, was a fallacy. He pointed out that two issues were confused in Klopfer's position. One was the personality as a whole governed by the laws which unit personalities follow and the second was the component elements, " . . . whether we are contemplating the psychological traits or the impersonal Rorschach factors which stand for these traits." Of the latter, Beck claimed that they can and must be isolated and subjected to experimentally controlled observation independently of the problems relating to the whole personality. "The criteria of the two are different and validation is within two totally different spheres of reference."

Point of View

As is obvious from this sampling, a position on the most fruitful approach to take in future Rorschach validation work is dependent upon one's theoretical and methodological biases in the broader problems within the fields of science and psychology. Accordingly, before embarking on a set of criticisms and suggestions, it is necessary in the interest of clarity to make an explicit statement of position on the crucial issues involved.

Validity is here considered to be measured by a statement of the degree of concomitant variation between two independent variables, one of which is designated as the criterion variable. The data obtained from a validity investigation indicate the degree of predictability of the magnitude of one variable, given the magnitude of the other. In the problem of Rorschach validation, then, the usefulness of the results will depend upon the criterion chosen. The requisites for an acceptable criterion variable are: (1) amenability to the assignment of numerals with a high degree of reliability which represents a relative statement of the degree of presence of the variable, and (2) known or hypothesized relationships between the criterion variable and other variables useful in work with personality. Use of the term "numerals" in (1) above indicates another bias. The writer believes that the variables of personality are lawfully related to relevant environmental variables and since laws are numerical statements for optimal serviceability, it is necessary to quantify the variables in order to find the laws. This is deemed a requirement for a science of psychology. Finding the laws depends in part upon a methodological approach geared to finding lawful relationships. It then becomes an empirical matter as to whether or not the lawfulness is "out there."

In the case of the Rorschach it is held as a working hypothesis that many of the variables of personality can be measured by it, the task being to ascertain the tenability of the hypothesis by making a series of verifiable hypotheses relating clearly defined personality variables to each of the factors or groups of factors found in a Rorschach protocol. It is also held that the laws describing the functioning of the whole personality are of the same class as those relating component parts to environmental variables. In describing the whole personality, what is needed is a composite law expressing the component parts and their interaction. Again, this may be true or false according to future empirical findings. Its immediate value lies in focusing on a search for relationships that can be expressed in a single system of laws. As a practical matter, it is probably more fruitful to start with component parts until more of the relevant variables for the total personality are ascertained. The rest of this paper will be devoted to an application of this orientation in an attempt to clarify some of the methodological problems inherent in Rorschach validation.

LEVELS OF INFERENCE

From a similar theoretical and methodological background, Steisel and Cohen (**34**) have elaborated on Beck's differentiation of the levels of inference used in interpreting from Rorschach protocols. To clarify the problems of Rorschach validation they differentiate three levels of inference, each of which require validation.

> Level I covers the inferences drawn from the raw response record for purposes of arriving at the appropriate scoring symbols and constructing an adequate psychogram.
>
> Level II represents the descriptive statements which can be made by inspection of the psychogram, sequence analysis and analysis of specific verbalizations.
>
> Level III includes the comprehensive, coherent personality diagnoses which may include the classification of the subject in some nosological group.

This can be considered a vertical schema ascending in order of degree of complexity (the number of variables involved in attempts to validate the inferences made). However, it can be pointed out that Level I does not involve the problem of validity as herein defined. Steisel and Cohen's Level I can more adequately be defined as entailing the problem of reliability, that is, the setting up of symbols each with a definition which summarizes certain aspects of a given response. The problem is to find the definition which will yield the highest degree of agreement in the case where several observers using the same definition for a given symbol are tested for the degree of concurrence of their applications of the symbol in question.

Having established that certain symbols can be applied with a satisfactory degree of reliability, Level II is reached. Level II represents the inferences as to specific processes or mechanisms which can be made about a personality. The raw data for these are found mostly in the psychograms (factors such as $F+\%$, M, C, $W:M$, etc.) though material such as symbolism, pathognomic signs, and sequence analysis (**19**) may also be used. As a point of departure some of the Rorschach studies sometimes referred to as validation studies in the literature will be classified in terms of the Rorschach variables used and the method employed. Each study will be evaluated in terms of the notion of validity expressed above.

Level II: Psychological Processes Represented by Psychogram Factors

1. *Hypnosis.* Sarbin (**31**) took Rorschach protocols from a single subject under four different conditions. Two of these conditions were with the hypnotic suggestion that the patient was Mme. Curie, and then Mae West; a third condition was that of hypnosis with no suggestion; and the fourth condition was in the waking state. He found that each of the three artificial situations produced changes in several of the psychogram factors and he attributed the changes to induced sets. However, this does not qualify as a validity study in that the relationship between the particular stimuli used to induce the sets, and changes resulting, are not known.

Levine, Grassi, and Gerson (**20**) used hypnosis in a Rorschach study with an improved method. They obtained a Rorschach protocol under normal conditions and then under several conditions of hypnotically induced affective changes. They suggested certain situations (under hypnosis) the psychological consequences of which were predicted on the basis of clinical experience. The Rorschach protocols obtained under the experimental conditions reflected changes in psychogram factors for each of the induced conditions. Further, the changes obtained in each case reflected the Rorschach factors usually associated with the respective affective conditions induced. Validity of the results in this case, is, of course, dependent upon the accuracy with which a group of judges can predict the psychological correlates of certain situations. Accurate prediction in such cases is dependent in part upon knowledge of the differential reactions of a large group of subjects to the same situation. Thus the relationship between the factors of the emotionally charged situation and the personality involved must be known in order to predict for small groups. Also, since it is known that hypnotizability is related to certain aspects of personality, results of such studies cannot legitimately be generalized beyond groups that can be hypnotized.

Bergmann, Graham, and Leavitt (**4**) used Rorschach protocols in conjunction with hypnotic age level regressions. The authors do not claim this to be a Rorschach validity study. They acknowledge using Rorschach personality pictures obtained at the various suggested age levels to validate the extent of regression, in that superficial regression would not be expected to reflect a complete change in Rorschach patterns. However, were the extent of regression under hypnosis established by use of independent measures other than Rorschach, then Rorschach factors could be validated by this procedure. Detailed histories of certain levels of individual development could be obtained and a group selected on the basis of optimum uniformity. On the basis of the personality features held in common for the individuals of this group, predictions could be made as to the relevant Rorschach factors that would be obtained under hypnotic regression to the respective levels of similarity. However, such a roundabout method is warranted only if there is a demonstrable need for an approach to the study of personality such as that used by Bergmann, Graham, and Leavitt.

In general, it might be said of use of hypnosis for Rorschach validation that considerably more data are needed to demonstrate how much of the personality changes are produced by the nature of the situation (i.e., effect of submission of the passive individual to the assertive hypnotist) and how much by the specific suggestions used.

2. *Drugs.* The studies reported in the literature (**13, 18, 27, 37, 38**) relating Rorschach factors to administration of drugs do not qualify as validation studies in that the laws describing the relationship between the drugs used (mescaline, histamine, amytal) and personality factors measured independently of the Rorschach are not known. The data

reported suggest that the reactions obtained depend in part upon features of the predrug personality. Guttman (**13**) states, "... there is need of a searching investigation in terms of the subject's setting and individual equipment before specificity of reaction can be established." Until the laws relating a single drug with personality reactions are found, this method will not be serviceable in Rorschach validation work.

3. *Operationally Defined Situations.* As Williams points out (**39**) this method requires: (1) setting up of an "independent and standardized criterion situation which would yield a quantitative index to an operationally defined aspect of personality, and (2) examination of the validity of those Rorschach factors primarily associated with this personality process in terms of the criterion measure." He assembled the definitions offered for intellectual control and designed a situation to measure this. He used as his measure decrement in performance of a task requiring intellectual factors under stress compared with performance under optimum conditions. Thus, the definition of intellectual control is in terms of the operations used to measure it. He administered a standard Rorschach to each of his subjects and used as the Rorschach index of intellectual control $F+\%$ and integration of form with color on the basis of hypotheses stated by Beck and Rapaport. He acknowledged the fact that other factors also were used in Rorschach interpretation and that he was operating at what is considered in this paper to be Level II. His results tend to bear out the hypothesized relationships between the selected Rorschach factors and the operationally defined aspects of personality. Structurally, this study conforms with the tenets presented here. However, he overlooked the fact that a linear relationship between intellectual control and $F+\%$ is not hypothesized by Beck and Rapaport for emotionally charged situations. An $F+\%$ above the optimum hypothetically represents an overcontrol due to inhibition and anxiety, compulsiveness, and rigidity or strong superego. Consequently, it might be hypothesized: (1) that overly-acute form perception is a defense against affective instability and as such might be related to loss of control under affective stress, and (2) that the threshold for affective stress would be lower than that for an individual with a lower but optimal $F+\%$. Therefore, it might be more in line with expectancy to adujst for $F+\%$ above the empirical mean for a standardization population, rather than consider it to be a continuous distribution with respect to intellectual control under stress.

Another study with a similar methodological basis provides a focus for the problem of the usefulness of the independent variable with which the Rorschach factors are correlated. Steisel (**33**) attempted to demonstrate a relationship between the factors on the Rorschach commonly hypothesized to be related to suggestibility, and performance in the autokinetic and Hull body sway situations. The relationship between suggestibility and performance in these situations is not as universally acceptable as is the relationship between intellectual control and the situation used by Williams. Therefore, acceptance of Steisel's results as bearing upon suggestibility as a personality variable depends upon the demonstration of a relationship between his two criterion measures and another independent measure of suggestibility. Insofar as his concept of suggestibility is operationally defined in terms of these two indices it is a defensible method at the current stage of development of Rorschach validation; i.e., if he had found a relationship between his selected variables it would then be a fruitful problem to experiment with the autokinetic phenomenon and Hull body sway technique to find their relationship to other independently defined aspects of personality. Even if they should be found related to aspects other than suggestibility, Steisel's data would be useful.

Rockwell, Welch, Kubis, and Fisichelli (**28**) report an experiment concerned with color shock. They define shock in terms of the neurological conception as a condition of lowered excitability. They measured shock by taking a continuous reading of the palmar skin response during Rorschach administration and described the changes expected with "shock" present and those expected with startle. Thus, their measure constitutes an operational criterion for shock and startle. They found that shock, as neurophysiologically conceived, does occur on those cards and in those subjects where it is expected, whereas startle does not. On the basis of these results they reject the Beck criteria for "color shock" which are based on expectation of a startle reaction to the color cards. This is a rather clear-cut example of an operationally defined situation devised to test two antithetical hypotheses. Acceptance of their interpretation depends upon evidence for relationship between the neurological conception of shock and Rorschach's conception of emotional inhibition. Assuming this relationship, it then remains to be determined what conditions the startle response (or more properly the Rorschach criteria for its presence) is related to—as a problem for those interpreters who use the "startle" criteria.

4. *Empirical Sign Studies.* The studies that fall into this group are concerned with demonstration of a relationship between certain single psychogram factors or groups of factors and effectiveness of various therapeutic techniques. (The sign studies concerned with relationship of Rorschach factors to psychiatric diagnostic categories will be dealt with in the series of Level III studies.)

In 1940 Piotrowski (**23**) published data which resulted from a study of two groups of schizophrenics, one of which showed improvement following insulin therapy while the other did not. He compared the pretreatment Rorschach protocols, and found that the presence of a response which included color as a determinant (*FC*, *CF*, or *C*) and for which there is a concrete or emotional association, i.e., other than color denomination, occurred significantly more frequently in the records of those who subsequently improved than those who did not. He reported that 81 per cent of the improved group met this criterion. Halpern (**14**) in a study of similar design found that the pretreatment Rorschach psychograms of those patients who subsequently improved with insulin therapy manifested greater productivity and those factors associated with "wider emotional range and greater capacity for empathy." Graham (**9**) with a study similarly designed found that those who improve with insulin give more responses determined by the chiaroscuro aspects of the inkblots. Benjamin (**3**) gives a preliminary report of a follow-up study of insulin-treated schizophrenics for whom he had pretreatment Rorschach protocols. Though his data are unquantified, he states that the preliminary suggestion is that the appearance of low $F+\%$, DdW and DW, and highly irregular sequence are associated with poor prognosis.

In these studies the criterion variable is judged improvement, and consequently judgment of improvement must be based upon clear-cut reliable criteria. Also, a clear statistical statement of the relationship between degrees of improvement and the signs reported to discriminate between "improved" and "unimproved" is required for an acceptable statement of validity. The above studies do not meet these criteria, but the results suggest possible relationships to be checked by more rigorous methods.

In an investigation concerned with factors related to success of metrazol therapy with a group of psychotics, Morris (**22**) improved on the method in the preceding studies. He provided for more clear-cut criteria for judgment of improvement and then selected critical points for his signs by referring his findings to the chi-square distribution for a statistical statement of his results.

The significance of this group of studies lies in the attempt to establish the usefulness of the Rorschach for prediction under restricted, specified conditions. However, they contribute little to validation of the Rorschach in terms of the second validity criterion, which calls for correlation between Rorschach variables and independent measures of personality for which a relationship with other personality measures is known. Should it later be demonstrated that improvement following insulin or metrazol administration is related to certain dimensions of personality, then the signs found in the above studies can be used as a basis for statements about these new dimensions insofar as the studies from which they result are experimentally sound. It should be pointed out that though these studies may not be of immediate optimal value for Rorschach validation, this does not mean they are equally questionable from the standpoint of other theoretical and/or practical considerations, such as the ones for which they were designed. They are on the same continuum only with respect to the first criterion of validity and are cited here merely to clarify the issues involved in Rorschach validation per se.

Level II: Psychological Processes Represented by Factors Other Than Those in the Psychogram

1. *Symbolism.* Earl (7) reported an experiment in which he investigated the validity of the usual interpretations for certain content categories frequently obtained in the Rorschach protocols of disturbed subjects. He was primarily interested in water responses and selected a group of feebleminded males aged thirteen to fifteen who were inmates of a residential school and had markedly immature personalities with severe conflict over masturbation. He selected those words with water content and all content words which occurred with the determinants *M*, *FM*, *m*, *K*, *FK*, and *KF* in the preexperiment protocols. He then put his subjects in an hypnotic trance and told them, "I am going to tell you to think of something. As soon as you think of it something quite different will come into your mind . . . I want you to say at once what it is. . . ." He then gave as stimulus words the words taken from the subject's own record which were suspected of being symbolic, interspersed among banal responses from his own record. He found that all the crucial words used for expressing a response the determinant of which was *m*, had symbolic significance. All responses having water as primary or secondary content "produced anxiety symbolized or overt in the hypnotized subjects." In interpreting his study he suggests that no common symbolic significance exists, at least among children, and that these responses, however determined, are associated with profound anxiety and it may be suspected that primitive sexual stirrings are always present.

The nature of his sample and small *N* (5 subjects) places serious restrictions on interpretations of these data. It would be desirable to have a control group of "stable" individuals so as to measure relative occurrence of water responses, and to compare the associations offered in the experimental situations by those "normal" cases in which water responses do occur with the responses of the experimental group. It might also be possible to do this type of work by use of word association without hypnosis, thereby freeing it from the limitations of hypnotic validation studies discussed above. By use of such a method, it is reasonable to expect that affective disturbances associated with the responses suspected of symbolic usage could be demonstrated. However, it is a more difficult problem to ascertain consistency for the particular material being symbolized by a given symbol. Demonstration of such consistency may require an approach which integrates data for responses with data on behavioral indices which are associated with aspects of the inkblots. An example of an attempt to associate a behavioral index with features of the blot is Blake's empirical study of ocular activity (5) during administration of the Rorschach technique. He found that number and duration of fixations was related to properties of the cards.

Card VI, for example, elicits the largest number of fixations for the top center detail. Card VI as a whole also elicits longer fixations during the initial inspection period than is true of the other cards in the sequence. It is also known that the top center detail of Card VI elicits the largest number of manifest sex responses and is suspected of producing symbolized sex responses as well as disturbances of sequence, form, control, etc., which are referred to as "sex shock." Therefore, it might prove fruitful to select those areas of the several cards which yield the highest number of fixations and use the responses offered for them in a method similar to that used by Earl. One might use a large group of normals and investigate the number of times a particular word is substituted in the word association experiment for the response offered during Rorschach administration.

2. *Empirical Sign Studies.* In 1941 Pitrowski (**24**) published a paper which presented six signs relevant to predicting the prognosis of schizophrenics given insulin therapy. It is unlike the similar studies reported above in that five of his signs are not found in the usual psychogram. He presented six signs which are reported to be indicators of a good prognosis: Generic Term, Variety, Evidence, Color Response, Indirect Color Approach, and Demurring. He offers what seems to be adequate criteria for determining presence of responses with these attributes, designates the method of determining improvement (psychiatric staff decision), and claims 93.3 per cent correct prediction by use of three or more signs as the criterion for good prognosis. As in the other studies of this type it remains to be demonstrated that prognosis for schizophrenics with insulin therapy is lawfully related to certain specified personality variables or processes. On the basis of his own data, Piotrowski offers an hypothesis in terms of intellectual regression to account for the differences between schizophrenics who improve with insulin and those who do not. Those who improve have undergone emotional regression; those who do not improve have undergone intellectual regression during their pretreatment illness. Assuming that there are operations with which to measure intellectual regression (other than Rorschach variables), this hypothesis then becomes subject to experimental verification. If verified, Piotrowski's signs would then become valid indicators of emotional regression (though the converse could not be assumed a priori in the case of absence of the signs). This is an illustration of how a sign study which, though not of itself a complete validation process, might lead to a series of studies the end result of which is validation of Rorschach variables. However, before results such as Piotrowski's are followed up, it would be advisable to select a new group of schizophrenics about to be treated with insulin shock and test the accuracy of the prediction of his signs. This would guard somewhat against the possibility that his original results were influenced by variables other than those to which he attributed his results.

Level III: Comprehensive Personality Diagnoses

At this level an attempt is made to relate variables of the Rorschach protocol to groups with various diagnostic labels. The emphasis is on a complete personality pattern rather than the component processes as far as the validity criterion is concerned. Thus one might measure the degree of relationship between diagnosis arrived at by use of the Rorschach and diagnoses made by any other means such as other clinical tests, psychiatric observation, a combination of tests, interviews and observations, or case history. Several such studies have been reported in the literature with results supposedly supporting the validity of the Rorschach as a diagnostic tool. However, the instability of current diagnostic groupings makes them unsuitable for use as validity criteria. Diagnostic practices differ from institution to institution and even from psychiatrist to psychiatrist in the same institution. A further criticism of this type of study can be made in terms of efficiency which would hold even if stability of diagnostic groupings were granted. The number of variables involved in arriving at a Rorschach diagnosis is large. Some of them might be highly related to the criterion and some not at all related. If the relationship of each variable (used in the Rorschach diagnosis) to the criterion variable is not known, then no statements can be made on possible improvements of the method of diagnosis. It is an all or none conclusion where there is ample reason for suspecting that some of the variables in combination might be more effectively used than others. Guirdham (**10**, **11**, **12**) has indicated one way to handle this problem in his work with epileptics and depressives in which he sought complexes of Rorschach variables each of which had some power to differentiate the groups in question from other nosological groups. He empirically tried numerical relationships between these variables until he found an index which provided optimum differentiation. The logical development of such an approach would lead to statistical analysis of a large number of Rorschach records from a variety of diagnostic groups, with each Rorschach variable being put into a separate multiple regression equation for each diagnostic category. Such an approach would constitute a sign method and, although a high order one, would still be subject to the limitations of sign approach as a validity method discussed earlier in this paper (i.e., the variables used for prediction are not selected on the basis of degree to which they enter into relationships with other variables of personality).

Piotrowski (**25**) has outlined three approaches to Rorschach diagnosis of mild forms of schizophrenia which can be applied to the general problems of Level III validation. He discusses first pathognomonic signs such as contamination, fluctuations in form level, and position responses. The limitations of this approach he lists as use of too small a part of the total record, and the observed fact that a large proportion of schizophrenics produce none of the signs, and thus absence of the signs does not mean schizophrenia is not present. He takes up next tabular diagnostic procedures which are based on Rorschach components that can be counted. It consists in establishing the differences in the frequency with which the scoring symbols occur in the records of schizophrenic and nonschizophrenic subjects. He indicates his preference for what he calls a systematic diagnostic procedure which "rests on principles which interrelate the components of the Rorschach method, expressing their mutual dependencies, and arranging them into a dynamic system." By this means he implies that the interrelated variables he uses are based on valid interrelationships among Rorschach symbols, e.g., that the more human percepts, the larger the percentage of sharply perceived forms.

His criticism of the pathognomonic and tabular procedures are well taken as far as he goes, but he overlooks the fact that the interrelationships used for the systematic procedure are not validated in the strict sense. Their "validation" rests on unquantified clinical observations. Further, the search for such relationships as used by Piotrowski requires a deduction as to component processes from what is known about the unit personality of schizophrenics. Such a search would be warranted only if the component processes for a given diagnosis were known. With regularly occurring relationships between given diagnostic categories and component processes we would have stable diagnostic categories, provided there were also operations to measure them. If these operations were independent of the Rorschach then there would be validity criteria for the Rorschach variables. However, in such an ideal situation Rorschach validation would then conform with the description for Level II validation.

CONCLUSIONS

In view of the foregoing considerations, the position is taken here that Rorschach validating procedures can be most fruitfully treated as problems in relating Rorschach variables to independent measures of component personality processes. There is no body of facts that can be invoked to prove this contention. It rests on principles of methodology concerned with the demonstration of valid relationships which are acceptable within a scientific framework. The merit of this approach applied to the Rorschach will be substantiated or discredited by the results obtained from its use. If procedures are used with provision for precise statement of results which are repeatedly obtained under the specified conditions, the Rorschach becomes a surer clinical device and it will facilitate psychological research. The clinician frequently protests that the experimental psychologist deals with artificial situations which are in no way related to the behavior which must be dealt with in the clinical problems of diagnosis and treatment. If the relationship between one or more Rorschach variables and component personality factors could be demonstrated, then those factors could be subjected to controlled experimentation with the Rorschach used to measure the changes that occur under varied conditions. It might also be possible to develop relationships between certain manipulable conditions and the origin of some of the personality components. Thus, a single measuring instrument, satisfactorily validated, could help to establish a connection between experimental results and clinical problems. It would seem particularly useful to aim toward this possibility in view of the widely held hypothesis that the Rorschach test samples a larger number of variables of personality than any other single instrument.

In conclusion it should be made explicit that the point of view on validation presented here does not do violence to the experienced interpreter's demands that each Rorschach factor be interpreted in terms of the other factors for any given record. What is required is specific statement of any particular constellation that is believed to be related to some measurable aspect of behavior, and then an appropriate test of this statement.

BIBLIOGRAPHY

1. Beck, S. J. Problems of further research in the Rorschach test. *Amer. J. Orthopsychiat.*, 1935, **5**, 100–115.
2. Beck, S. J. Error, symbol and method in the Rorschach test. *J. abnorm. soc. Psychol.*, 1942, **37**, 83–103.
3. Benjamin, J. D. A method for distinguishing and evaluating formal thinking disorders in schizophrenia. In J. S. Kasanin (Ed.), *Language and thought in schizophrenia*. Berkeley and Los Angeles: Univ. of California Press, 1946.
4. Bergmann, M. S., Graham, H., & Leavitt, H. C. Rorschach exploration of consecutive hypnotic age level regressions. *Psychosom. Med.*, 1947, **9**, 20–28.
5. Blake, R. R. Ocular activity during administration of the Rorschach test. *J. clin. Psychol.*, 1948, **4**, 159–170.
6. Brosin, H. W., & Fromm, Erika O. Some principles of Gestalt psychology in the Rorschach experiment. *Rorschach Res. Exch.*, 1942, **6**, 1–15.
7. Earl, C. J. C. A note on the validity of certain Rorschach symbols. *Rorschach Res. Exch.*, 1941, **5**, 51–61.
8. Frank, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, **8**, 389–413.
9. Graham, Virginia T. Psychological studies of hypoglycemia therapy. *J. Psychol.*, 1940, **10**, 327–358.
10. Guirdham, A. The Rorschach test in epileptics. *J. ment. Sci.*, 1935, **81**, 870–893.
11. Guirdham, A. Simple psychological data in melancholia. *J. ment. Sci.*, 1936, **82**, 649–653.
12. Guirdham, A. The diagnosis of depression by the Rorschach test. *Brit. J. med. Psychol.*, 1936, **16**, 130–145.
13. Guttman, E. Artificial psychoses produced by mescaline. *J. ment. Sci.*, 1936, **82**, 203–221.
14. Halpern, Florence. Rorschach interpretation of the personality struc-

ture of schizophrenics who benefit from insulin therapy. *Psychiat. Quart.*, 1940, 14, 826–833.

15. Harrison, R. The thematic apperception and Rorschach methods of personality investigation in clinical practice. *J. Psychol.*, 1943, **15**, 49–74.
16. Hertz, Marguerite R. Rorschach: twenty years after. *Psychol. Bull.*, 1942, **49**, 529–572.
17. Hertz, Marguerite R. The Rorschach method: science or mystery. *J. consult. Psychol.*, 1943, **7**, 67–80.
18. Kelley, D. M., Levine, K., Pemberton, W., & Lillian, K. K. Intravenous sodium amytal medication as an aid to the Rorschach method. *Psychiat. Quart.*, 1941, **15**, 68–73.
19. Klopfer, B., & Kelley, D. *The Rorschach technique.* New York: World Book Co., 1942.
20. Levine, K. N., Grassi, J. R., & Gerson, M. J. Hypnotically induced mood changes in the verbal and graphic Rorschach: a case study. *Rorschach Res. Exch.*, 1943, **7**, 130–144.
21. Macfarlane, Jean W. Problems of validation inherent in projective methods. *Amer. J. Orthopsychiat.*, 1942, **12**, 405–410.
22. Morris, W. W. Prognostic possibilities of the Rorschach method in metrazol therapy. *Amer. J. Psychiat.*, 1943, **100**, 222–230.
23. Piotrowski, Z. A. A simple experimental device for the prediction of outcome of insulin therapy in schizophrenia. *Psychiat. Quart.*, 1940, **14**, 267–273.
24. Piotrowski, Z. A. The Rorschach method as a prognostic aid in the insulin shock treatment of schizophrenics. *Psychiat. Quart.*, 1941, **15**, 807–822.
25. Piotrowski, Z. A. Experimental psychological diagnosis of mild forms of schizophrenia. *Rorschach Res. Exch.*, 1945, **9**, 189–200.
26. Rapaport, D. In Rorschach forum report. *Rorschach Res. Exch.*, 1939, **3**, 107–110.
27. Robb, R. W., Kovitz, B., & Rapaport, D. Histamine in the treatment of psychosis. *Amer. J. Psychiat.*, 1940, **97**, 601–610.
28. Rockwell, P. V., Welch, L., Kubis, J., & Fisichelli, V. Changes in palmar skin resistance during the Rorschach test. I. Color shock and psychoneurotic reactions. *Monthly Rev. Psychiat. Neurol.*, 1947, **113**, 9–152.
29. Rorschach, H. *Psychodiagnostics.* (Transl. by P. Lemkau & B. Kronenberg.) New York: Grune and Stratton (distr.), 1942.
30. Rosenzweig, S. Outline of a cooperative project for validating the Rorschach test. *Amer. J. Orthopsychiat.*, 1935, **5**, 121–123.
31. Sarbin, T. R. Rorschach patterns under hypnosis. *Amer. J. Orthopsychiat.*, 1939, **9**, 315–318.
32. Sargent, Helen D. Projective methods: their origins, theory and applications in personality research. *Psychol. Bull.*, 1945, **42**, 257–293.
33. Steisel, I. An experimental investigation of the relationships between some measures of the Rorschach test and certain measures of suggestibility. Unpublished doctoral dissertation, Univ. Iowa, 1949.
34. Steisel, I., & Cohen, B. D. The problem of validation of the Rorschach test with special reference to the method of direct experimentation. Unpublished paper, 1947.
35. Thurstone, L. L. The Rorschach in psychological science. *J. abnorm. soc. Psychol.*, 1948, **43**, 471–475.
36. Varvel, W. A. Suggestions toward the experimental validation of the Rorschach test. *Bull. Menninger Clin.*, 1937, **1**, 220–226.

37. Wertham, F., & Bleuler, M. Inconstancy of the formal structure of the personality; experimental study of the influence of mescaline on the Rorschach test. *Arch. Neurol. Psychiat.*, 1932, 28, 52–70.

38. Wilkins, W. L., & Adams, A. J. The use of the Rorschach test under sodium amytal and under hypnosis in military psychiatry. *J. gen. Psychol.*, 1947, 36, 131–138.

39. Williams, M. An experimental study of intellectual control under stress and associated factors. *J. consult. Psychol.*, 1947, 11, 21–29.

Received February 8, 1950.

CRITIQUE OF HORNELL HART'S "SOCIAL SCIENCE AND THE ATOMIC CRISIS"[1]

RALPH H. GUNDLACH
New York, N. Y.

Recognition of the importance of war, peace, and social change as research problems for social scientists has been expressed in the $1000 bond award which Mr. Bernays has asked SPSSI to administer. A distinguished committee of social scientists chose a work of Professor Hornell Hart for the award in 1948, and the summary of his book was published as a supplement to the *Journal of Social Issues*. Both author and editor expressed the hope that the thinking of Prof. Hart would provoke discussion and further planning.

Hart's underlying propositions seem sound and well documented: that over the centuries the power to destroy has increased at an accelerating rate while the power to control such forces—although increasing—lags so far behind that destruction may overcome our civilization.

It is Hart's proposed solution to this problem which strikes me not only as wishful thinking but as dangerously misleading. The fact that his formulation is selected as the best contribution to the problem by so distinguished a committee leads me to believe that my objections to Hart's proposals are not shared by many. I shall formulate my objections, then, as sharply and clearly as I can.

Hart's analysis and program carry, I believe, three false and misleading assumptions which are so fundamental as to vitiate his program, and perhaps lead in directions opposite to those stated.

1. Hart seems to assume that knowledge, reason and equalitarian justice, implemented by persuasion, are the ruling factors in social organization; and that self-interest and the possession of an advantage in power and/or ownership are of negligible importance. He says, for instance, that depressions and war are examples of cultural lag which can be caught up "only [by] a swift upsurge in the effectiveness of social science" (p. 5). "Cultural evolution produces . . . a wide variety of culture complexes which are not adequately integrated. . . . Two outstanding examples of this tendency are economic depression and war" (p. 10). "At present social science is radically inadequate for grappling with the atomic crisis" (p. 14). "The crux of the atomic crisis is the lag of social science behind the accelerating upsurge in the technologies of

---

[1] HART, HORNELL. Social science and the atomic crisis. *J. soc. Issues*, Supplement Series No. 2., April, 1949. A condensation of the contribution which won the Edward L. Bernays Atomic Energy Award of the Society for the Psychological Study of Social Issues.

destruction" (p. 17). Thus, according to Hart, the ills of the world, including wars and depressions, are due to simple lag; and to bring about world peace and prosperity he proposes his prize-winning project in social science research.

But reality stands to the contrary. What retains wars and depressions is not the absence of knowledge, but institutionalized "vested" interests. Labor-management disputes are settled largely by tests of strength. Slavery was not given up as a consequence of an educational program on race differences; on the contrary, a mythology of race was formulated and promulgated at home and abroad to justify and perpetuate discriminatory practices.

Although Hart's program ignores the power-structure of society, it is obvious that his proposed band of social science experts, working out the program for social change, social control, etc., abandon the whole structure and ideology of the "free enterprise system," and move in the direction of some kind of technocratic, planned economy. Not that I hold any brief for competitive or monopolistic capitalism, but Hart's program must quickly reach the power-arrangements of the status quo, and go with it or against it. Money and resources are more readily available to programs that extend and tighten the established controls than to programs that call for relaxing autocratic controls and spreading power and benefits more widely. Hart's program, by ignoring the central power relationships, tacitly accepts and endorses the status quo—even perhaps bids for its economic support. This suggests the possibility of its being subsidized to become an instrument of resistance to social change.

2. The second and third "wrong assumptions" may be deduced from the following quotations, which are part of a questionnaire Hart submitted to a group of students in a preliminary study concerning "the relative desirability of various possible outcomes of the Atomic Crisis" (p. 22). He asked them "three hypothetical questions":

> 1. If we had to choose between giving Russia a free hand, and engaging in a fairly evenly balanced atomic war with her, which would do the least damage in the long run to the basic values of civilization?
>
> 2. If there were one chance in ten that World War III could be avoided for the next 25 years by working peaceably through the United Nations with all its defects, and four chances in ten that Russia could be rendered helpless by the scientific use of bacteriological and radioactive sanctions, before she could damage America seriously, which course of action would be most likely to preserve the basic values of civilization?
>
> 3. If there were five chances in ten that the United States could maintain law and order in the world by independently expanding her naval and air forces, and by keeping far in advance of other countries in military research, and two chances in ten that destruction could be avoided by creating a cooperative federation of the democracies and pooling all our military resources

> together, which policy would be most likely to preserve our basic values? (p. 22).

I wish to emphasize two assumptions from these quotations. First, it is assumed that the people of the United States, their government, and their major dominant policy-making groups, are unified, and that this unity includes agreement on altruistic, world-wide goals, and that peace is high among them. The other assumption, reflected in the phrase "preserve our basic values" and "values of civilization," implies in this context that the American values, as expressed by the current policies of the United States Government, are morally right and are identical with the basic values of civilization. In parallel it assumes that the policies and acts of the political leaders of other states, insofar as they differ from United States policy (especially, that is, Russia) are evil, wrong, and destructive to "civilization's values."

Such a view ignores the important body of facts, readily available for instance, regarding power politics (**1**, **5**); world relations toward Soviet Russia (**4**, **5**); our own State Department's distortions of historical events regarding the Soviet Union (**6**); the expanding post-war American imperialism (**2**); and the post-war effects of the general policy upon our fundamental democratic rights (**3**).

Hart's questionnaire follows the naive ethnocentric identification of "we-with-God" versus "they-with-the-Devil" for the parties set up in the conflict; and with moral self-righteousness, modestly proposes for serious consideration a war of extermination against the Soviet Union as one way to peace and the expanding values of civilization. I do not see how this implies the use of accelerated social science to overcome the growth and extension of destructive power.

A point of view similar to Hart's is expressed in simple terms of economic and military mastery by Dr. Virgil Jordan, President of the National Industrial Conference Board, in these terms:

> Let us first offer the utmost capacity of our economic power for reconstruction freely to every people who will undertake to abolish all national military expenditures and disarm down to the level of the local constabulary.
>
> Let us, secondly, demand the unlimited right of continuous inspection and control of every industrial operation and process and every public policy which may have the most remote relationship to armament and warfare.
>
> And finally, let us make, improve and keep plenty of our best and biggest atomic bombs for that imperative purpose: let us suspend them in principle over every place in the world where we have any reason to suspect evasion or conspiracy against this purpose; and let us drop them in fact, promptly and without compunction, where ever it is defied (**2**, pp. 175–176).

I hope I am wrong, but I find nothing in the assumptions which underlie Prof. Hart's proposals toward solving the Atomic Crisis that would disagree with the position of Dr. Jordan, who suggests the military and economic domination of the globe by the group of which he was president.

BIBLIOGRAPHY

1. Brady, R. A. *Business as a system of power.* New York: Columbia Univ. Press, 1943.
2. Marion, G. *Bases and empire: a chart of American expansion.* New York: Fairplay Publishers, 1948.
3. Rogge, O. J. *Our vanishing civil liberties.* New York: Gaer Associates, 1949.
4. Sayers, M., & Kahn, A. E. *Great conspiracy: the secret war against Soviet Russia.* Boston: Little, Brown, 1946.
5. Schuman, F. L. *International politics: an introduction to the western states system.* New York: McGraw-Hill, 1937.
6. Schuman, F. L. *Introduction to "Falsifiers of History: an historical document on the origins of World War II."* New York: Committee for Promotion of Peace, April, 1948.

Received March 24, 1950.

REJOINDER TO GUNDLACH'S CRITIQUE

HORNELL HART
Duke University

Dr. Gundlach's first objection appears to be summarized in his statement: "What retains wars and depressions is not the absence of knowledge, but institutionalized 'vested' interests." He fears that a research program which does not "call for relaxing autocratic controls and spreading power and benefits more widely" may be "subsidized to become an instrument of resistance to social change." His position appears to be in line with the familiar dogmas of Socialists and Communists. One of the major tasks of social science (as I see it) should be to accept as hypotheses such propositions as those quoted above from Dr. Gundlach, to test them impartially, to reach whatever verifiable conclusions may be obtainable, and to develop inductively a comprehensive, operational, scientific theory of social change, in which should be fitted whatever of truth may be found to lie in the socialistic, as well as in other beliefs.

Even if Dr. Gundlach's position were entirely sound, the fact would remain that the capitalistic holders of power, as well as the politicians, the university faculties, the newspaper industries, and the others whom Socialists and Communists regard as henchmen of the capitalistic class, all stand in imminent and crucial peril—not so much of losing their *relative* advantage over their fellow human beings as of seeing the collapse of the entire structure of civilization upon whose continuance their present power and advantage—and even their very existence—depend. If there is such a common danger confronting all members of civilization, it would certainly seem to be a proper function of social science to seek to clarify as operationally as possible the causal factors in this crisis, and to discover what possible lines of action might have the greatest promise of averting the catastrophe and of keeping civilization in operation.

What effects a genuinely scientific study of the total situation would have upon the distribution of power, wealth, and privilege is another question. As I see it, the most wealthy and powerful people in the world today are, to very considerable extents, defeating their own most fundamental needs and potentialities. But whether or not that is true, I do not see how a genuine social scientist can merely accept the communistic dogmas about human motivation, and use them as axioms, without thereby abdicating as a scientist.

The remainder of Dr. Gundlach's paper is concerned mainly with the question of whether the values of civilization depend upon the victory of the Western democracies over Soviet Russia. It can be agreed at once that both the Soviet Union and the Western powers have at times broken their promises, have used misleading and even false propaganda, have infringed the rights of weaker peoples, have used military force and other forms of coercion when just solutions might have been reached by agreement, and have been guilty of other violations of democratic ideals. It can also be agreed (I believe) that both the Russians, the Americans, and the other peoples of the world desire deeply to preserve their cities from being bombed, to improve rather than degrade their present economic standards of living, to increase the opportunities of their youth for education, to decrease disease and eliminate preventable deaths, to cultivate the arts, to encourage joyous recreation among their people, and the like. The Soviet Communists seem to believe that these things can be accomplished only by universalizing the dictatorship of the proletariat. This system has developed out of centuries of history, during which the Russian people have had practically no direct experience with parliamentary government, civil liberties, or economic and political freedom, but have been constantly subjected to secret police, repressive censorship, and underground conspiracies.

In sharp historical and factual contrast, the United States is one of twelve nations which have demonstrated, during the past century and a third, that it is possible for democracies to work out their international relations to their mutual propserity and security without war against each other. These twelve countries include the five English-speaking nations, the three Scandinavian countries, and France, Switzerland, Netherlands, and Belgium. Each of these twelve nations has a constitution which recognizes the people, acting through their elected representatives, as the supreme authority. Their democracy is the product of centuries of experiment, experience, and evolution. Elections in these countries are competitive, with freedom to organize rival political parties and to seek to influence voters and public opinion through a free press, radio, and public discussion (distorted though these often are by financial and other pressures). Civil liberties are established constitutionally, and in spite of many serious violations are on the whole sustained.

Governments in these countries are avowedly for the public welfare. "Left-wingers" claim that the rights of private property and the interests of the wealthy have obten been maintained at the expense of human welfare; business men complain that labor unions are allowed to dictate and that the rich "get soaked" through taxation. The civil and economic rights of Negroes are often violated—particularly in the states of the Deep South. But probably no people in all history have progressed so rapidly toward economic freedom, good housing, education, and health as have the Negroes in the United States.

All of these countries have shared in the accelerating growth of science, interchanging freely their inventions and discoveries and the applications of these to industry, agriculture, health, amusements, and other interests—particularly to military destructiveness. All have developed a swiftly growing body of trained scientific technicians and experts. Science, like government, industry, and religion, has been strongly competitive in nature but, as in these other fields, organized cooperation is increasingly taking the place of individualistic competition.

The standards of living of city workers and of farmers are higher in these twelve countries than anywhere else in the world and have been increasing more rapidly than in other nations. Death rates are lower, and expectations of life higher, than anywhere else.

Numerous factual studies of Communistic countries report that the individual citizen is regarded by the Soviet government as having no rights except as he submits to, and identifies himself with, the "party line" established by the self-appointed and self-perpetauting leaders of the proletarian dictatorship. Individual and mass murder, wholesale sabotage, plundering of occupied countries, exploitation of slave labor, complete suppression of all discussions not approved by the dictatorship—such policies are avowed or are taken for granted as instruments of the revolution. Under the democratic conception of social science, such charges as these are matters which call for impartial factual investigation—a task which has not yet been performed adequately by American social scientists. But under the Communistic control, no social scientist is permitted to hold a position, or to teach or carry on research, unless he dogmatically assumes and proclaims the immense superiority of all Communistic programs and actions over those of the non-Communistic democracies.

A major assignment for which social scientists are already vigorously laying foundations is a definitive study of the social psychology of Soviet Communism. The layman who wants to understand Russia finds himself confronted with an enormous mass of literature. Much of it is passionate propaganda. Much more is sincere but distorted arguments by fanatical Communists and anti-Communists. Other large portions of this literature are superficial, more or less sensational, poorly grounded in factual knowledge. Buried amidst this misleading and inaccurate body of writings are numerous painstaking, dispassionate, searching studies of available facts about the history, culture, psychological attitudes, resources, ideas, and plans of the Russian people and of the Soviet government. We need a sifting of information and conclusions about the Soviet Union. We need a clear consensus of the agreements shared by the best-informed authorities, and a clear setting forth of the major issues which exist between such authorities.

Dr. Gundlach challenges the capacity of social scientists to exert constructive and disinterested influence upon social policies. My own analysis on this point would be as follows. Social action is the result of applying social power. Vast reservoirs of power exist, consisting of such factors as the number of men of military age, the gold reserve of a country, its natural resources, its factories, buildings, machines, and other productive equipment, its atomic fuels, its educated brain-power, the loyalties of its people to the government, to political parties, to labor unions, and the like. The problem is to utilize and restrain these energies in such ways as to preserve the vital structures of civilization, and also to serve to the greatest practicable degree the basic values of mankind. Since those basic values include freedom, the problem involves the attainment of the necessary control by means of consent, with the minimum necessary amounts of control by prophylactic coercion.

The uses to which national and world power are put, and the types of control employed, are determined by the directors of power, consisting of that minority which legislates, and which determines the selection of political and economic executives. The directors of energy, in turn, are held in check or released by the few great triggers of power, such as the Rockefellers, Henry Ford, the President of the United States, the chairmen of certain Congressional committees, the Secretary of State, the Supreme Court, John L. Lewis, and the like.

One of the basic tasks of social science is to understand thoroughly, dispassionately, and predictably just how these reservoirs, directors, and triggers of power operate. Hindsight keeps telling us how bunglingly these powers have been applied in the past. The survival of civilization depends upon a really adequate operational analysis of the way that social forces work, and the methods by means of which the best brains of America and of the world can team up to redirect human power into constructive—or at least not disastrously destructive—lines of action. In the future, social power may gravitate increasingly into the hands of those who assimilate and make use of a truly effective science of social relations.

Received July 25, 1950.

ANDERSON, JOHN E. *The psychology of development and personal adjustment.* New York: Henry Holt, 1949. Pp. v+720. $3.25.

This book was designed for a general course in Human Development and Personal Adjustment in the two-year curriculum of the General College of the University of Minnesota. In contrast to a typical course in general psychology, this volume emphasizes "the genetic approach and treats adjustment in greater detail."

A little over half of the book presents concise accounts of development in such areas as physical growth, skills (sensori-motor, language, problem solving, intelligence), motivation, emotion, attitude, interest, and social behavior. The second half analyzes the adjustment process and then attacks problems of adjustment in college, in maturity and old age, in vocation, and in family life.

The most impressive thing about this excellent text lies in the way in which the author (1) develops for the student the value of an objective, constructive approach to personal problems and (2) demonstrates again and again that such an approach is possible. He does this by consistently showing that he regards the student as an adult, by selecting important issues and interesting, useful information, and by maintaining without effort professional integrity in the analysis and discussion of very difficult problems.

Here and there the particular psychological concepts with which Anderson structures his text seem somewhat irrelevant to the general aim of the book. This reader wonders whether, for example, the Lewinian concepts presented in the chapter on Motivation can contribute, without considerably more development, to the goal of "understanding at the terminal level."

The book can be highly recommended to the teacher of any introductory course which emphasizes developmental concepts and personal adjustment.

VINCENT NOWLIS.

State University of Iowa.

GUTHRIE, EDWIN R., & EDWARDS, ALLEN L. *Psychology: a first course in human behavior.* New York: Harper, 1949. Pp. x+315. $3.00.

Students should enjoy this book. Its distinguishing characteristic is that it starts on the student's common-sense side of each topic and then relates the findings of psychology to that every-day experience. This strategic emphasis is implemented by Guthrie's well-known luminous style, filled with apt examples. The Flesch Ratio is close to 4, which indicates that the *potential* audience has finished eighth grade or one year of high school while the *typical* audience has finished high school or one year of college. According to Flesch, college students read material of this degree of difficulty for fun.

The choice of contents, too, was apparently made with the student reader in mind. Of eighteen chapters, the first three are introductory, giving something of the development of modern psychology and introducing and discussing some aspects of science. The other two emphases of the book are on learning (including remembering), which has four chapters (68 pages), and on various aspects of personality and adjustment, which have four chapters (70 pages).

On the other hand, certain topics were apparently thought to be more appropriate for advanced courses and are treated more briefly. A 15-page chapter on "The Structural Basis of Behavior" covers material on muscles, glands, the structure of receptors, the nervous system and brain function. Sensory and perceptual phenomena in the various sense-modes are described in a 26-page chapter on "Attending and Perceiving." Descriptive statistics along with reliability, validity, and standard scores are nicely introduced and summarized on pp. 165–169. The chapter on the topics of "Heredity and Maturation" has 13 pages. The relative amount of space devoted to the other four chapters on Emotion, Drives, Intelligence, and Thinking is about the same as in other introductory texts.

In the introductory chapters, some aspects of the scientific point of view are especially well presented. The authors recognize that objectivity depends upon agreement (p. 18). They are clear about scientific determinism: "One of the assumptions that science must make is that all events are determined by antecedent conditions. This is not a belief . . . ; it is an assumption" (p. 27). And they know that science is concerned with "prediction and control" (p. 31) instead of the halfway stations of *understanding* and *explanation*.

Yet with this emphasis on the scientific point of view, there is no mention of the structure of an experiment, no mention of independent and dependent variables, nor even of "controls." The student is given no aid in constructing an experiment for himself. To the reviewer this omission is regrettable, although it may have been made in the interests of simplicity and clarity. In line with this possible plan of the authors, relatively few experiments are completely described. Experimental *results* and *conclusions* are given on almost every page, but rarely in such a way that the student reader can see how the procedure and plan lead to the conclusions. There are few tables and only 17 figures.

The treatment of *learning* presents the senior author's associative theory: In a particular situation, the organism learns the *last act performed before* the behavior sequence is interrupted by reward. To the reviewer, this description of learning differs only slightly from theories which say that any act will be learned if *followed immediately by reward.*

With only small modification the treatment in this book will be acceptable to advocates of most current theories.

RICHARD P. YOUTZ.

Barnard College, Columbia University.

V. HOLZSCHUHER, LUDWIG. *Praktische Psychologie; die Primitivperson im Menschen.* (Practical psychology; the primitive person in man.) Seebruck: Heering-Verlag, 1949. Pp. 763.

As a result of the decline of psychology in Germany, few if any significant German contributions have come to the attention of American psychologists in recent years. For this reason, if for no other, it may prove illuminating to take a look at a rather comprehensive work which purports to deal with "practical psychology" as the German psychologist sees it. Holzschuher's *Praktische Psychologie* is a post-war contribution and presumably a representative synthesis of the German viewpoint in psychology. Its value to American readers is considerably enhanced by the fact that the author shows a familiarity with developments in American psychology, some of which he severely criticizes. As a specialist in advertising, his treatment of the effects of advertising on the public's purchasing behavior, public opinion, and attitudes is most expert. His approach is essentially "practical" but the reader familiar with German psychological literature will not be surprised at the notable lack of experimental data. Thus, the orientation is not primarily *geisteswissenschaftlich*, although large sections of the book fall into this somewhat ill-reputed category.

The book is divided into two parts, a theoretical section (about 100 pages), and a major second part, which deals with the application of the author's theoretical framework to his subject matter (about 590 pages). The latter contains chapters on the influence of unconscious factors upon various aspects of human behavior.

Holzschuher's conceptual scheme, which he calls a "strata theory," is a variety of depth psychology. Throughout he is imbued with the realization, amply substantiated by analytic psychology but not always seriously considered in academic circles, that the supremacy of the ego in directing human behavior is largely illusory. He postulates an *Ichperson* (ego) and a *Primitivperson* (id) and refers to the tensions characterizing the interaction of these systems as "bipolar." There is nothing intrinsically original about these formulations, although the author belabors the differences between his system and the theories of Freud and Jung. Nonetheless, we are presented with a consistent application of a theoretical framework to a variety of psychological phenomena, such as the behavior of social groups in response to propaganda aimed at the primitive layers of the individual personality, and people's reactions to advertising campaigns. American advertising provides source material for the author's thesis of unconscious, primitive determinants and their effect upon behavior. It may be mentioned here that the book is handsomely illustrated, in particular with pictures from the American scene.

The fact that the ego is constantly concerned with rationalizing unconscious tendencies is of far-reaching significance in psychological experimentation. American investigations of public opinion polls and attitude measurements are criticized for their failure to take this into account. The author asserts that they usually yield merely contents of ego consciousness which are only conditionally identical with reality; for this reason it is difficult to distinguish rational from rationalized material in such procedures. In this connection, the following statement, in the reviewer's opinion, is well worth quoting: "The work of institutes of public opinion research . . . is therefore to be considered with the greatest caution, at least so long as they continue to ignore the most elementary insights of modern depth psychology" (p. 152 f.).

Students of propaganda will find many interesting first-hand observations and analyses of Nazi propaganda whose effects upon the masses are interpreted by Holzschuher in terms of its appeal to the individuals' primitive impulses. The following example is quite instructive:

> All those who rationally rejected [Nazi ideology] must admit that they were sometimes seized by a frenzy of enthusiasm which penetrated their veins and ran down their backs when they became eye and ear witnesses of well organized demonstrations. Here human bipolarity really reveals itself: The ego says "stupid trash," the primitive person jubilates (p. 245).

The treatment of personality and the emotions is studded with a wealth of anecdotal material and generalizations which are difficult to evaluate. The positive contribution of the work seems to lie chiefly in the author's penetrating analysis of Nazi ideology and propaganda, and its disastrous influence upon mass behavior. For this reason alone it deserves careful study by students of propaganda and social psychologists in general.

Human Resources Research Laboratories,
Department of the Air Force.
HANS H. STRUPP.

THORNDIKE, EDWARD L. *Selected writings from a connectionist's psychology.* New York: Appleton-Century-Crofts, 1949. Pp. vii+370. $3.50.

In the preface to *Selected Writings from a Connectionist's Psychology,* Thorndike states that the volume was the result of a request for some three hundred pages of his most important contributions to psychology.

The book is admittedly not that, however, because he omitted from selection much of his earlier work including that on animal intelligence, individual differences, intercorrelations of abilities, the application of the scientific method to problems of learning and teaching, and also his later studies of individual differences in communities, and all of his work on tests. He favored (1) materials which are not now readily available for class reference because printed in journals or monographs and (2) books likely to go out of print. In making the selections from these sources, he states, "I have, in fact, had in mind especially the needs of teachers who wish their students to know something of connectionist psychology at first hand. . . ."

Between the Preface and Chapter I, which begins the contributions, he gives his autobiography taken from *A History of Psychology in Autobiography* (Murchison, Vol. III, 1936), supplemented by two new pages covering the period from 1934 to 1948. To the reviewer this section provides the most fascinating reading of the entire volume. In it one can discern as truly as in the scientific writings the bent of the man. As responsiveness to outer pressure or opportunity rather than to inner need is revealed as the chief source of problems of study, so also is the more outstanding personal characteristic of striving to find out what things or processes are like by experiment rather than by discussing what someone else has said they are like.

The selections are arranged in twenty-two chapters beginning with the "Law of Effect" and ending with "Darwin's Contribution to Psychology." In between are covered subjects almost as extensive as psychology itself. Several chapters deal with highly technical problems involving variables bearing on the establishment of *connections*, while others treat problems of a more general nature such as personality, labor, mental fatigue, language, and values. The forty-odd selections were taken from eight books, seventeen scientific periodicals, one popular magazine, and three "memorial" volumes. The overall range in publication dates was from 1905 to 1947, with two-thirds of the articles being originally published after 1931. Lengths of reprints vary from three to eighteen pages with the average around nine.

While it will be possible for a student to learn from *Selected Writings* much about Thorndike's connectionism after 1930 and less about his connectionism before 1930, the more thorough investigator will not be content with the brevity and sketchiness of many of the reprints. The extensive topical coverage and the lack of coherence of the chapters indicate, however, that the author did not intend the book for that purpose. Its greatest usefulness will be in the hands of students not specializing in the research aspects of learning and motivation but rather in their applications.

JOHN B. WOLFE.

University of Mississippi.

WEVER, ERNEST GLEN. *Theory of hearing.* New York: John Wiley, 1949. Pp. xiii+484. $6.00.

Professor Wever's book is addressed to the difficult and challenging problem of evaluating the so-called "place," "frequency," and "place-frequency" theories of hearing from a historical point of view and in the light of modern anatomical, neurological, and psychological evidence. This is a major contribution from one of the principal investigators in the field of hearing, culminating over two decades of continuous study and research on the subject. It is written in a straightforward and lucid style.

Part One of *Theory of Hearing* sets forth the classical and present day theories of hearing with critical insight into their shortcomings. Here more than a score of theories responsible for stimulating and guiding thought and research in audition are arranged and discussed with distinguished scholarship.

On this firm historical basis, Part Two takes up the modern developments of classical theories. The impact of new knowledge of the detailed structure and dynamics of the cochlea and the electrophysiology of the auditory nerve upon the classical Helmholtzian theory of sound analysis is demonstrated. In this same section the "volley" principle of nerve action is presented as a phenomenon which must, it is argued, supplement and to some extent supplant the theory of localization on the basilar membrane as a mechanism whereby the auditory system achieves an analysis of frequency; the author feels that the latter theory has gained an undeserved popularity. Wever takes the position that the principles of "place" and "frequency" should not be considered necessarily as conflicting and that they can be harmoniously combined to achieve a fairly comprehensive and workable theory of hearing. Although considerable attention is given to attributes of audition other than pitch discrimination, this theme dominates the book.

In Part Three, which constitutes over half the book, the author presents his "resonance-volley" or "place-volley" theory of audition. "Resonance" in the physical sense of sympathetic vibration of any cochlear structure is discounted, but variation of place of stimulation on the basilar membrane as a function of the frequency of a sound wave is held to account for discrimination among tones of high frequency (above about 4000 cps); the analysis of low frequency sounds is made possible, it is proposed, at a cortical level due to a somewhat faithful reproduction of the wave form of sound stimulation by a patterning of auditory nerve impulses according to the principles of nerve "volley" activity. This proposed dual mechanism is based on well-documented evidence that the conditions necessary for frequency localization within the cochlea are met and, likewise, it is established that to some degree the auditory nerve discharges at rates that correspond to the frequency of sinusoidal sound stimulation up to near 4000 or 5000 cps.

The experimental evidence on acoustic neurology, sensitivity, loudness and fatigue, pitch and pitch discrimination, auditory abnormalties, tonal interaction, temporal phenomena of auditory perception, and binaural phenomena are marshalled before the reader in an orderly fashion and cast against the "place-volley" theory for explanation. Original data hitherto unpublished will be found here. Finally, the book ends with a section of definitions and symbols and a select bibliography of about five hundred titles.

It is to be expected that there are those who will disagree with the interpretation and treatment given some of the auditory data and feel that insufficient facts are available concerning cochlear function to warrant the emphasis placed upon the proposed role played by volley action within the auditory nerve. However, all students and investigators of hearing will find *Theory of Hearing* an important book as a source of information and stimulation that will have a significant influence upon future auditory research.

K. D. Kryter.

Human Resources Research Laboratories,
Department of the Air Force.

Rubin, Edgar. *Experimenta psychologica.* (Psychological experiments.) Copenhagen: Ejnar Munksgaard, 1949. Pp. 356. 3100 kroner.

"Friends and students have expressed their interest in the publication in collected form of the scientific papers which I have published in various periodicals. This has led to the present book." Thus does the author state the purpose of this reprint of fifteen of his articles published between 1911 and 1938. Eleven of them are in German, two in English, and two in French. Rubin speaks of his own as "adspective" psychology which is the opposite of structuralism and yet not quite like the orthodox Gestalt psychology to whose early development he has contributed so significantly with the figure-ground concept. Adspective psychology is concerned with aspects of wholes which are not to be artificially analyzed and subsequently synthesized. One does not know how many aspects there are or can be. Classification of these aspects is always tentative, as new material will often make regrouping desirable. Such a point of view runs counter to the modern tendency toward systematization in psychology.

Almost one half of the volume, 170 pages, is devoted to the single problem of how a straight edge is experienced by vision and touch. Other problems treated by the author are the latency between stimulation of the retina and awareness of the stimulus, the visual perception of actual movements, an auditory complication experiment, and time, vibratory, temperature, and taste sensations. A few short articles are concerned with theoretical questions, and one deals with an aspect of the philosophy of Hoefding, his teacher.

The total impression gained from this volume is that of the life work of a psychologist who was a lucid writer, a careful experimenter, and an independent thinker who early in his career went his own way. It is probably safe to predict that Rubin's lasting reputation will not rest on the contributions which are here reprinted, but on his *Visuell wahrgenommene Figuren.*

KARL F. MUENZINGER.

University of Colorado.

CROW, LESTER D., & CROW, ALICE. *Educational psychology.* New York: American Book Co., 1948. Pp. xiv+599. $4.00.

The unusually detailed and systematically organized table of contents of this text well displays the range and arrangement of topics touched upon. The 29 chapters are grouped into seven parts. After two introductory chapters, Part II considers Growth and Development, with separate brief chapters given to the general nature of growth, and to physical, mental, emotional, and social development. Part II on "The Learner's Potentialities" deals successively in chapters 8–11 with intelligence, aptitudes, personality traits, and individual differences. Six chapters then treat learning: general principles, attitudes, study habits, motor skills, thinking, and transfer. Part V in three chapters considers the nature, scope, and techniques of evaluation, and statistics. Six chapters on the "Psychology of Learning Areas" are devoted to the following topics: "developing skill in spoken language" and in reading, spelling, handwriting, foreign language; the "Psychology of Arithmetic" and of algebra and geometry; the natural sciences in the elementary school and in the high school, and similarly the social sciences at both levels; the development of appreciation in art, music, and literature; educating for health and safety; and "Vocational Exploration and Preparation." A concluding part on adjustment considers exceptional individuals, behavior drives, and mental hygiene. Obviously, the authors have attempted to cover an exceedingly wide range of topics. The dominant emphasis, as stated in the preface, is upon the guidance of development toward good adjustment.

With so many and various topics to be dealt with, inadequacy in the treatment of at least some would seem hardly avoidable. The course of physical growth from birth to maturity is presented in 9 pages, with two charts. The psychology of reading, including remedial teaching, is considered in 5½ pages, plus two plates showing an ophthalmograph and records obtained therewith. Spelling and handwriting receive each three pages, social sciences in elementary and secondary school a total of 4 pages, "vocational exploration and preparation" 7 pages. As is perhaps inevitable when so much is attempted, large use is made of secondary sources and some rather old material. Treatment may become superficial, even platitudinous (thus we are told that "Sufficient restful sleep is healthful" in the two pages on "physical well-being" in the chapter on mental hygiene). Over-caution in brief statements may omit and seem to reject present best-informed attitudes on a subject; thus a one-page discussion of "the sex urge" could almost have been given to a Sunday school class in 1890—though later comments on "sex adjustment" suggest a more tolerant understanding.

The volume is, however, a well organized display of the variety of topics in the total broad field of educational psychology, and is written with an evident sincerity of interest in children and their welfare. Each one of the 27 chapters ends with "questions and topics for discussion" and a list of "selected references" as aids to instructor and student. The author and subject indices seem more carefully made than in many similar volumes.

SIDNEY L. PRESSEY.

Ohio State University.

ROSENZWEIG, SAUL (with KATE L. KOGAN). *Psychodiagnosis.* New York: Grune and Stratton, 1949. Pp. x+380. $5.00.

In this book the authors present an introduction to methods of psychological diagnosis of individual personalities. Their aim has been to provide a variety of concrete illustrations which exemplify the conceptual and clinical operations actually performed by the modern, or dynamically oriented clinical psychologist. The authors de-emphasized abstract formulations of theory and method "partly because it has seemed premature to state baldly principles that are as yet largely implicit and tentative in the work of the psychodiagnostician" (p. viii). Abstract and theoretical formulations have by no means been disregarded, however, in the attempt to avoid jargon.

The three major sections of this book introduce the reader to diagnostic methods by demonstrating, in steps of increasing complexity and scope, the manner in which the working clinician uses and integrates test results and other information. First, under the heading of "The Range of Available Instruments," typical and popular tests for the assessment of intelligence, deterioration, vocational aptitudes and interests, and personality are described and illustrated by protocols and discussion. The next section, "The Process of Psychodiagnostic Integration," presents test results and protocols for four different cases. The discussion of each case examines the contribution of particular tests to the diagnosis, as well as the proper combination and correlation of test results to provide an integrated picture of each personality. This section also includes a nice discussion of the art and practice of writing psychodiagnostic reports. In the section, "A Variety of Applications," the authors present additional cases to give the process of diagnostic integration more extended treatment. This section emphasizes in greater detail the contribution of social and historical data to the understanding of the total individual.

The general introduction, "Clinical Psychology as a Psychodiagnostic Art," presents the basic frames of reference which influence the authors throughout the book. Here the authors distinguish between the more old-fashioned psychometric approach to diagnosis, which they seem to consider as depending largely upon rigid and stereotyped inference from tables of norms, and more modern approaches. The latter emphasize highly individualized interpretations which rely heavily upon inferences from qualitative aspects of objective test performance and projective test results, all in accordance with dynamic theories of personality. This does not mean that the authors deny the importance of conventional norms but rather that they emphasize, in a quite specific fashion, the relative and provisional contribution such normative data make to the understanding of an individual person.

In the closing chapter on "Psychodiagnosis as a Science" the authors view with favor the beginning rapprochement between experimental and clinical psychology. Here they make a case for the individual, in all his complexity, rather than the "generalized average man" (*sic*) as the more profitable unit for purposes of scientific scrutiny. The discussion in this chapter seems almost to identify scientific enterprises devoted to the discovery of general laws of behavior with the production of propositions about "the generalized average man" or about isolated segments of behavior which will be irrelevant to the task of understanding the total human individual. Though such an outcome has been frequent, possibly even modal, this reviewer is inclined to question any implication that it is either necessary or unavoidable. After all, psychoanalytic theory, which is widely conceded to be relevant to the clinical enterprise of understanding the individual, is a general theory of human behavior and development. And recent advances in learning theory, based largely upon isolated segments of the behavior of the rat, appear not only to confirm some aspects of psychoanalytic theory, but even promise to furnish clues as to how it may be revised and extended.

In general, the discussions as well as the test and case materials are well formulated and lucidly expounded. They should prove effective as an introduction to the sophisticated clinical approach to psychological diagnosis. The authors deserve commendation for their attack on this difficult educational task. Though a text of this length cannot cover, or even mention, all of the popular testing techniques, the selection in this book seems sufficiently broad and representative for introductory purposes.

For all its virtues, the book has several serious limitations in its role as an introductory text. For one thing, the interpretations and inferences included in the discussions range from the probably valid or widely accepted and canonical, through the hypothetical, to those which the authors would, no doubt, be willing to class as frankly speculative. Even so, alternative interpretations and qualifying statements appear only infrequently. The beginner, without extensive background in clinical psychology, will find it difficult to evaluate the validity of these interpretations and inferences. Little in the book would even suggest that such a problem exists. Furthermore, the bibliography includes only very scanty reference to the critical research and clinical literature. Though space limitations imposed on such a book preclude detailed examination of the validity of the interpretations presented, the general problem deserves more emphasis than it received. For pedagogical reasons some attempt should have been made to differentiate between standard, widely accepted interpretations and those which are hypothetical. This could have been achieved, without swamping or discouraging the beginner, by means of brief, organized discussions illustrated by some research data and argument, all presented with the same lucidity the authors achieved in delineating their approach to psychodiagnosis.

HOWARD F. HUNT.

University of Chicago.

STECKLE, LYNDE C. *Problems of human adjustment.* New York: Harper, 1949. Pp. x+351. $3.00.

This text is offered by the publishers for courses in mental hygiene, personal adjustment, and personality development. The author states that the volume represents the concepts he has found to work when the attempt is made to introduce man to himself.

It suffers grievously from the lack of a systematic conceptual framework, and the net effect is chaotic. It is evident, for example, that the author has not sought for essence in the field of motives. Contradictions abound. While he urges acceptance of man's animal nature, which he identifies with emotion, he seems unable to see emotion as anything but bad. His casualness with this concept as with others creates trouble here as in many other places. Emotion and learning are flatly called antithetical. It is said that under heightened emotion one learns with difficulty. Yet fear and anger as habits are maintained because they assist the person in the attainment of his goals.

One fears that the young undergraduate would be led to think that the best he could do with his own emotions would be to sit on them. Emotions like delight and affection are mentioned, but briefly, and they are given slight roles in conduct. It will be hard to convince the undergraduate, too, that "the biological act of intercourse is a not too important part of what the term love has come to imply."

The author maps "The Road to Neurosis" in a chapter, and it is of interest that nowhere does the road traverse the domain of Eros. Eros does not even go on the journey as guide or companion, nor is he granted the honor of being a symbol of that security and self-esteem which all the travelers, although strayed from the proper road, are held to be seeking.

Although his world sounds in many ways like a comfortable one where the main problems of life are really pseudo-problems which "do not exist for the well-adjusted individual," the author develops a degree of hortatory zeal in urging a rational course upon man, and a vehemence in castigating him for his stubborn irrationality, that would please the soul of many a one of the orthodox preachers who are among the groups most vigorously thrashed. Rules, guides, and precepts for realistic and efficient living are liberally set forth, although the author repeatedly professes complete disbelief in the efficacy of telling man what he should do.

Students, undergraduate or other, will thrive better on more profound, more deeply searching treatises than this one.

DWIGHT W. MILES.

Western Reserve University.

McKINNEY, FRED. *The psychology of personal adjustment.* (2nd Ed.) New York: John Wiley, 1949. Pp. xi+752. $6.00.

Freshmen in college are in position to profit from a knowledge of psychology before taking general psychology, usually not given until the sophomore year. Aware of this need, McKinney in 1941 wrote the first edition of the present volume which, in his words, was "written to meet the need for a basic psychological text which frankly attacks the problems of the student that are most vital to his personal adjustment and which offers him factual material on these problems." The content of the second volume in general is identical with the first, with relatively few changes. The addition of a chapter on development of personality and the addition of more recent references are the most noticeable changes.

There are a number of striking features in the book. One is a large amount of illustrative case material; another is a comprehensive list of references; still another is the inclusion of many lists which offer in summary fashion advice or maxims for living. The book includes lists which give a college student's reasons for going to college; a check list for effective study habits; one giving motives for the individual; and another a list of issues for a personal philosophy of habits which make us likeable; advantages and disadvantages of fraternity life; and offensive and inoffensive ways of handling different situations. Advice is offered on ways of taking exams, budgeting time and many diet and health habits, developing a personal philosophy, developing leadership, and controlling masturbation. Counseling is favored because discussion provides an outlet, de-sensitizes the individual, changes his attitudes, discovers repressed material, helps an individual organize and develop self-acceptance and motivation. All of these are described in brief as assets and potentials of counseling. The dynamics of the process by which such virtues are obtained are barely mentioned. In the chapter devoted to the need for a philosophy of life, Mencken, Keith, Hu Shih, Lilienthal, and Mumford are quoted, and fairly effective use is made of these quotations for suggesting a need for attaining a philosophy of life.

The emphasis on growth and maturity is not as dynamically presented in the material considered as the students obtain from McKinney's courses. He belongs to that group of good teachers who influence by radiation as well as by presentation. There is material in this book which can be used to advantage by students who are on the search for knowledge about life and even for gaining some insights which will make it possible for them to get a look at life in a manner that is forward-looking, healthy-minded, and integrated rather than merely adjusted—though adjustment, not integration and integrity, is the emphasis of the book. All in all, though, it can serve very well the purpose for which it is intended—a helpful volume for the person who is starting on the road towards leading a college life which promises well for his growth.

H. MELTZER.

Psychological Service Center, St. Louis.

ALMY, MILLIE CORINNE. *Children's experiences prior to first grade and success in beginning reading.* New York: Bureau of Publications, Teach. Coll., Columbia Univ., 1949. Pp. viii+124. $2.35.

GRAY, WILLIAM S. *On their own in reading.* New York: Scott, Foresman, 1948. Pp. xvi+268. $2.00.

WITTY, PAUL. *Reading in modern education.* Boston: Heath, 1949. Pp. xvi+319. $3.50.

These three books cover phases of reading, a topic which has been widely discussed in professional literature for the last few decades. One of the books, *Children's Experiences Prior to First Grade and Success in Beginning Reading*, is an experimental study. The other two, *Reading in Modern Education* and *On Their Own in Reading*, are discursive presentations aimed at helping the teacher to do a better job in the teaching of reading. They could both be used as textbooks in university course work.

Reading in Modern Education covers the whole field of reading from the development of readiness in the preschool years to at least a mention of reading problems of college students. The last chapter is a summary statement concerning the author's own philosophy of the importance of reading and its status in the future.

Children's Experiences Prior to First Grade and Success in Beginning Reading is an attempt, on the part of the author, to determine experimentally whether the factors which have been considered to be important in the development of readiness actually do produce the effects desired. As far as experimental design is concerned it is somewhat questionable whether the author has been able to prove her point, but it is good to review these factors and call them again to the attention of those who read the book. The difficulty is that, cast in this style, it probably will have very little reading public. Yet written for the potentially more effective public—parents—it might receive wide attention. The author concludes that "This study has substantiated the fact that there is a positive relationship between the child's responses to reading opportunities prior to the first grade and his success in learning to read (p. 114). . . . The correlation would never approach the point where beginning reading success could be predicted on the basis of preschool reading experiences. There are too many other contributing factors" (p. 117). The author goes on to state that even if exact prediction cannot be made, it is important to know that this relationship does exist.

On Their Own in Reading is an illustration of a book written to a purpose and satisfying that purpose very fully; that is, to aid teachers in the development of lesson plans and specific classroom exercises which will bring about more effective classroom instruction in word attack and thus develop more independence in reading. This topic is mentioned and given some attention by Witty in his discussion of reading in general, but Gray has not only told the teacher what to do, he has shown her how to do it. This is something which is so often not effectively done. In education today we have a tendency to talk at the students instead of talking with them and demonstrating ways in which topics covered by our lectures may be carried out, thus taking the first step toward helping the teacher to put them into practice.

There is little doubt that the teacher who takes the trouble to read *On Their Own in Reading* will at least try some of the techniques so ably presented. A sound psychological approach to the teaching of reading is well illustrated. For instance, psychologists recognize the soundness of the philosophy that the background of the individual is a potent factor in learning. Dr. Gray does not let his readers forget this. He gives illustrations where this principle might otherwise be overlooked. Thus, there is the reminder that the teacher must start teaching with words in the spoken vocabulary of the child for, he says, at no time does the child get material solely from printed symbols. He must at all times bring meaning to them. In much reading instruction and formal course work, we have a tendency to forget these basic principles and make presentations appear to stand alone, forgetting that the background of the individual affects markedly what he obtains from the material presented.

While the other two books cover only restricted phases of the field, Witty leaves almost no topic untouched in his presentation. It takes both kinds of books to present a working approach to the subject. Witty's somewhat cursory presentation of some topics will serve as at least an introduction. He has presented bibliographical references which will help the reader to go to more detailed presentations such as those by Almy, Gray, and others to get the details which would be necessary if the topics discussed by Witty were to be well understood. At times his presentation is little more than a full outline of his subject. This is a good illustration of where his reader must bring a good deal to the topic presented if this book is to be fully understood. It is well to mention that Witty ably makes a balanced, concise presentation of such topics as the "non-verbal" child, "dominance," the use of mechanical gadgets in the teaching of reading, and other such subjects which have for too long caught the fancy of some persons in the field. They have built theories of the psychology of reading around these restricted considerations and thus base their instruction on somewhat narrow concepts rather than on a total over-all sound theory of what is known of the psychology of reading. Witty's very brief section on reading as it is related to the total psychological development of the child is also well done and should be of special interest to both parents and teachers.

FRANCES ORALIND TRIGGS.

University of Maryland.

BOOKS AND MATERIALS RECEIVED

BARAHAL, GEORGE D. *Converting a veterans guidance center.* Stanford, Calif.: Stanford Univ. Press, 1950. Pp. viii+99. $1.50.

BARTLEY, S. HOWARD. *Beginning experimental psychology.* New York: McGraw-Hill, 1950. Pp. viii+483. $4.00.

BETTLEHEIM, BRUNO. *Love is not enough.* Glencoe, Ill.: Free Press, 1950. Pp. ix+386. $4.50.

BLOOM, BENJAMIN S., AND BRODER, LOIS J. *Problem solving processes of college students.* Chicago: Univ. of Chicago Press, 1950. Pp. 109. $2.75.

BORING, E. G. *A history of experimental psychology.* (2nd Ed.) New York: Appleton-Century-Crofts, 1950. Pp. xxi+777. $6.00.

BROWN, WARNER, AND GILHOUSEN, H. C. *College psychology.* New York: Prentice-Hall, 1950. Pp. 507. $3.75.

BURKS, BARBARA, AND ROE, ANNE. *Studies of identical twins reared apart. Psychological Monographs* No. 300 (Vol. 63, No. 5). Washington, D. C.: American Psychological Assn., 1949. Pp. v+62. $1.00.

CATTELL, RAYMOND B. *Personality: A systematic, theoretical and factual study.* New York: McGraw-Hill, 1950. Pp. xii+689. $5.50.

CRAFTS, L. W., SCHNEIRLA, T. C., ROBINSON, ELSA E., AND GILBERT, R. W. *Recent experiments in psychology.* (2nd Ed.) New York: McGraw-Hill, 1950. Pp. xvii+503. $3.50.

DASHIELL, J. F. *Student's Manual to accompany Dashiell's "Fundamentals of general psychology."* New York: Houghton Mifflin, 1950. Pp. 130. $1.00.

DEMING, WILLIAM E. *Some theory of sampling.* New York: Wiley, 1950. Pp. xvii+602. $9.00.

DERBER, MILTON. (Ed.) *Proceedings of the second annual meeting.* Publication No. 4. Champaign, Ill.: Industrial Relations Research Association, 1949. Pp. viii+299.

Edwards, Allen L. *Experimental design in psychological research.* New York: Rinehart, 1950. Pp. xiv+446. $5.00.

Elonen, Anna S. *A comparison of two tests of intelligence administered to adults. Psychological Monographs* No. 306 (Vol. 63, No. 11). Washington, D. C.: American Psychological Assn., 1949. Pp. 35. $1.00.

Festinger, Leon, Schacter, Stanley, and Back, Kurt. *Social pressures in informal groups.* New York: Harper, 1950. Pp. x+240. $3.00.

Fisher, Seymour. *Patterns of personality rigidity and some of their determinants. Psychological Monographs* No. 307 (Vol. 64, No. 1). Washington, D. C.: American Psychological Assn., 1950. Pp. v+ 48. $1.00.

Fisher. V. E. *The meaning and practice of psychotherapy.* New York: Macmillan, 1950. Pp. xv+411. $5.00.

Fryer, D. H., and Henry, E. R. (Eds.) *Handbook of applied psychology.* New York: Rinehart, 1950. Pp. xix+842. $12.50.

Greenblatt, Milton, Arnot, Robert, and Solomon, Harry C. *Studies in lobotomy.* New York: Grune & Stratton, 1950. Pp. x+495. $10.00.

Hahn, Milton E., and MacLean, Malcolm S. *General clinical counseling.* New York: McGraw-Hill, 1950. Pp. xi+375. $3.50.

Harris, J. Donald. *Some relations between vision and audition.* Springfield, Ill.: Charles C Thomas, 1950. Pp. v+56. $1.50.

Hartmann, A. A. *An experimental examination of the thematic apperception technique in clinical diagnosis. Psychological Monographs* No. 303 (Vol. 63, No. 8). Washington, D. C.: American Psychological Assn., 1949. Pp. v+48. $1.00.

Heron, William T. *Clinical applications of suggestion and hypnosis.* Springfield, Ill.: Charles C Thomas, 1950. Pp. ix+116. $3.00.

Honkauara, Sylvi. *On the psychology of artistic enjoyment.* Helsinki: Soumalainen Tiedaekatemia, 1949. Pp. 161.

Hunt, J. McV., Blenkner, Margaret, and Kogan, L. S. *Testing results in social casework.* New York: Family Service Assn. of America, 1950. Pp. 64. $2.00.

Hunt, J. McV., and Kogan, L. S. *Measuring results in social casework.* New York: Family Service Assn. of America, 1950. Pp. 79. $1.50.

Jennings, Helen Hall. *Leadership and isolation.* (2nd Ed.) New York: Longmans, Green, 1950. Pp. xvii+349. $4.00.

Kasius, Cora (Ed.). *A comparison of diagnostic and functional casework concepts.* New York: Family Service Assn. of New York, 1950. Pp. 169. $2.00.

Kates, Solis L. *Rorschach responses related to vocational interests and job satisfaction.* *Psychological Monographs* No. 309 (Vol. 64, No. 3). Washington, D. C.: American Psychological Assn., 1950. Pp. v+34. $1.00.

Koskenniemi, Matti. *Comments on group work in school.* Helsinki: Suomalainen Tiedeakatemia, 1950. Pp. 44.

Landis, Carney, and Bolles, M. Marjorie. *Textbook of abnormal psychology.* (Rev. Ed.) New York: Macmillan, 1950. Pp. x+634. $5.00.

LeFever, D. W., Turrell, A. M., and Weitzel, H. I. *Principles and techniques of guidance.* (Rev. Ed.) New York: Ronald Press, 1950. Pp. xvi+577. $4.25.

Mackworth, N. H. *Researches on the measurement of human performance.* Medical Research Council Special Report Series No. 268. London, Eng.: His Majesty's Stationery Office, 1950. Pp. 156. 6/6.

Menninger, Karl A., and Devereux, George. *A guide to psychiatric books with suggested basic reading list.* New York: Grune & Stratton, 1950. Pp. vii+148. $3.50.

Merry, Frieda K., and Merry, Ralph V. *The first two decades of life.* (2nd Ed.) New York: Harper, 1950. Pp. xiii+600. $3.75.

Morgan, Clifford T., and Stellar, Eliot. *Physiological psychology.* (2nd Ed.) New York: McGraw-Hill, 1950. Pp. ix+609. $5.00.

Munn, Norman L. *Handbook of psychological research on the rat.* New York: Houghton Mifflin, 1950. Pp. xxvi+598. $7.50.

Moreno, J. L., and Enneis, J. M. *Hypnodrama and psychodrama.* New York: Beacon House, 1950. Pp. 56. $2.75.

Parten, Mildred. *Surveys, polls, and samples: practical procedures.* New York: Harper, 1950. Pp. vi+194. $4.00.

Porteus, Stanley D. *The Porteus Maze Test and intelligence.* Palo Alto: Pacific Books, 1950. Pp. vi+194. $4.00.

Pressey, Sidney L. *Educational acceleration: Appraisals and basic problems.* Bureau of Educational Research Monographs No. 31. Columbus, Ohio: Ohio State Univ. Press, 1950. Pp. xii+153. $2.50.

Read, Katherine H. *The nursery school.* Philadelphia: Saunders, 1950. Pp. viii+264. $3.50.

Sargent, S. Stansfeld. *Social psychology.* New York: Ronald Press, 1950. Pp. x+519. $4.50.

Sharif, M. M. *Beauty and expression.* Pakistan: Muhammad Ashraf, 1947. Pp. 133. 7/6.

Sharif, M. M. *Beauty—objective or subjective?* Pakistan: Muhammad Ashraf, 1947. Pp. 87. 4s.

Sharif, M. M. *The nature of tragedy.* Pakistan: Muhammad Ashraf, 1947. Pp. 110. 7/6.

SKINNER, CHARLES E. (Ed.) *Elementary educational psychology.* (2nd Ed.) New York: Prentice-Hall, 1950. Pp. xv+592. $3.75.

STOTT, D. H. *Delinquency and human nature.* Fife: Carnegie United Kingdom Trust, 1950. Pp. viii+460.

STOUFFER, S. A., GUTTMAN, LOUIS, SUCHMAN, E. A., LAZARSFELD, P. F., STAR, S. A., AND CLAUSEN, J. A. *Measurement and prediction.* (Vol. IV of *The American Soldier.*) Princeton: Princeton Univ. Press, 1950. Pp. x+756. $10.00.

SULLIVAN, ALBERT J., AND MCKELL, THOMAS E. *Personality in peptic ulcer.* Springfield, Ill.: Charles C Thomas, 1950. Pp. x+100. $3.00.

TROW, WILLIAM C. *Educational psychology.* (2nd Ed.) New York: Houghton Mifflin, 1950. Pp. ix+761. $4.00.

ULETT, GEORGE. *Rorschach introductory manual.* St. Louis: Educational Publishers, 1950. Pp. 44. $3.00.

VON BONIN, GERHARDT. *Essay on the cerebral cortex.* Springfield, Ill.: Charles C Thomas, 1950. Pp. xiii+150. $3.75.

WELLS, CHARLES A. *The value of an oral reading test for diagnosis of the reading difficulties of college freshmen of low academic performance. Psychological Monographs* No. 308 (Vol. 64, No. 2). Washington, D. C.: American Psychological Assn., 1950. Pp. iv+35. $1.00.

WIENER, NORBERT, *The human use of human beings.* New York: Houghton Mifflin, 1950. Pp. 241. $3.00.

WILDE, OSCAR. *De Profundis.* (Ed. by Vyvyan Holland.) New York: Philosophical Library, 1950. Pp. 147. $3.00.

WITKIN, H. A. *Perception of body position and of the position of the visual field. Psychological Monographs* No. 302 (Vol. 63, No. 7). Washington, D. C.: American Psychological Assn., 1949. Pp. iii+46. $1.00.

Combat psychiatry. U. S. Army Medical Department, Vol. IX, Supplemental number. Washington, D. C.: U. S. Gov't Printing Office, 1949. Pp. ix+216. $.60.

Report of the Medical Council for the Years 1945–1948. Committee of Privy Council for Medical Research. London, Eng.: His Majesty's Stationery Office, 1949. Pp. 283. 5s.

INDEX OF SUBJECTS

Analysis, factorial, of animal behavior, 235
Analysis of variance, qualification in the use of, 130
Animal behavior, factorial analysis of, 235
Audio-visual aids for teaching of clinical psychology, survey of, 137

Children's drawings, studies in the psychology of, 1928–1949, 369
Chi-square test, misuse of, reply to Lewis and Burke, 331
    use and misuse of, comments on Lewis and Burke article, 338
    use and misuse of, 2×2 contingency table, 341
    use and misuse of, further discussion of, 347
Clinical psychology, teaching of, survey of audio-visual aids for, 137
Curve fitting, empirical, use of method of successive differences in, 260

Economics, psychological problems of, 289
Effect, circularity of law of, 52
Empirical curve fitting, use of the method of successive differences in, 260

Factorial analysis of animal behavior, 235
Fitting of Poisson and negative binomial distributions to observed accident data, validity of inferences drawn from, 434
Four studies in psychology and social status, 263

Galvanic skin response, nature of, 97

Hart, Hornell, critique of his "Social science and the atomic crisis," 509

Latin square design, homogeneity of variance and, 118
Leadership, membership and organization, 1
London's historical survey of psychology in the Soviet Union, note on, 146
    reply to note on, 150

Membership, leadership and organization, 1

Negative binomial and Poisson distributions, fitting of to observed accident data, 434

Organization, leadership and membership, 1

Pain perception, recent research on, 449
Poisson and negative binomial distributions, fitting of to observed accident data, 434
Psychodiagnostics, present-day, statistical techniques in, 475
Psychology in the Soviet Union, London's historical survey of, note on, 146
    reply to note on, 150
Psychology and social status, four studies in, 263

Reinforcement, partial, review and critique, 193
Repression, experimental analogue of, historical summary, 39
Rorschach validation, methodological aspects, 493

Skin response, galvanic, nature of, 97
Social attitudes, indirect assessment of, 15
Social psychology, what constitutes a good textbook in, 152
"Social science and the atomic crisis," critique of Hornell Hart's, 509
    Hart's rejoinder to Gundlach's critique, 513
Social status, psychology and, four studies in, 263
Soviet Union, London's historical survey of psychology in, note on, 146
    reply to note on, 150

Successive differences, use of method of,
  in empirical curve fitting, 260

"Use and misuse of the chi-square test,"
  comments on, 338

2×2 contingency table, 341

Variance, analysis of, qualification in the use of, 130
  homogeneity of, and the Latin square design, 118

INDEX OF AUTHORS

GENERAL ARTICLES, NOTES, SPECIAL REVIEWS

Burke, C. J., 347
Campbell, D. T., 15
Centers, R., 263
Edwards, A. L., 118, 341
Edwards, W., 449
Garner, W. R., 260
Goodenough, F. L., 369
Gundlach, R. H., 509
Harris, D. B., 369
Hart, H., 513
Hayes, S. P., Jr., 289
Jenkins, W. O., 193
Katz, E., 137
Lemmon, V. W., 130
Lewis, D., 347
London, I. D., 150
Maritz, J. S., 434
McCleary, R. A., 97
Meehl, P. E., 52
Mensh, I. N., 475
Park, D. G., 152
Pastore, N., 338
Peters, C. C., 331
Razran, G., 146
Royce, J. R., 235
Schneider, L. I., 493
Stanley, J. C., Jr., 193
Stogdill, R. M., 1
Webb, W. B., 130
Zeller, A. F., 39

BOOKS REVIEWED

Alexander, F., 189
Almy, M. C., 529
Anderson, G. L., 279
Anderson, J. E., 517
Andrews, T. G., 166
Appicciafuoco, R., 172
Bell, J. E., 89
Britt, S. H., 162
Brouwer, P. J., 362
Buros, O. K., 177
Chase, S., 78
Combs, A. W., 170
Coombs, C. H., 181
Crow, A., 524
Crow, L. D., 524
Crutchfield, R. S., 155
Dashiell, J. F., 356
Dennis, W., 273
Deri, S. K., 360
Donahue, W. T., 181
Dunbar, F., 189
Edwards, A. L., 517
Egan, J. P., 272
Escalona, S. K., 185
Farnsworth, P. R., 163
Frank, J., 284
Frank, L. K., 89
French, T. M., 189
Gray, W. S., 529
Guthrie, E. R., 517
Harris, N. G., 93
Humphrey, G., 277
Hurlock, E. B., 183
Hutchinson, E. D., 277
Johnson, D. M., 81
Johnson, P. O., 285
Kanner, L., 187
Katz, B., 87
Kogan, K. L., 525
Korner, A. F., 361
Krech, D. C., 155
LaPiere, R. T., 163
Lief, A., 185
Lindesmith, A. R., 282
Machover, K., 84
McKinney, F., 528
Mathewson, R. H., 180
Mayo, E., 85
Mead, M., 280
Murphy, G., 275
Mursell, J. L., 365
O'Kelly, L. I., 357
Parry, J. B., 363
Pollak, O., 164
Postman, L., 272
Rosenzweig, S., 525

Rubin, E., 523
Ryan, T. H., 174
Salter, A., 358
Schaefer-Simmern, H., 165
Sheldon, W. H., 448
Smith, G., 80
Snygg, D., 170
Stacey, C., 279
Stagner, R., 446
Steckle, L. C., 527
Strauss, A. L., 282
Super, D. E., 178
Swenson, E. J., 279
Symonds, P. M., 86
Thorndike, E. L., 520
Thorndike, R. L., 76
Thorpe, L. P., 87
Travers, R. M. W., 181
Underwood, B. J., 444
Valentine, W. L., 274
Vernon, P. E., 363
v. Holzschuher, L., 519
v. Schiller, P., 83
Wever, E. G., 522
Wickens, D. D., 274
Witty, P., 529
Wolff, W., 91
Zipf, G. K., 172

BOOK REVIEWERS

Anastasi, A., 80
Anderson, E. E., 81
Bayley, N., 361
Billingslea, F. Y., 91
Britt, S. H., 284
Corsini, R. J., 172
Crafts, L. W., 185
Davis, F. B., 365
Edgerton, H. A., 285
Ericksen, S. C., 274
Finan, J. L., 444
Finger, F. W., 272
Fischer, R. P., 183
Fletcher, F. M., Jr., 180
Fryer, D. H., 363
Gregory, W. S., 87
Guetzkow, H., 277
Hahn, M., 178
Heathers, G., 282
Hunt, H. F., 525
Hunt, W. A., 448
Jenness, A., 163
Johnson, D. M., 275
Johnson, H. M., 174
Kay, L. W., 84
Kendler, H. H., 279
Kryter, K. D., 522
Louttit, C. M., 170
McKinney, F., 187
Meltzer, H., 528
Metcalf, J. T., 164
Meyer, H. J., 78
Miller, J. G., 85
Miles, D. W., 527
Mowrer, O. H., 358
Mueller, C. G., 166
Muenzinger, K. F., 523
Munn, N. L., 273
Nowlis, V., 517
Pattie, F. A., 83
Peak, H., 172, 356
Postman, L., 155
Pressey, S. L., 524
Riess, B. F., 189
Sanford, F. H., 446
Sarason, S. B., 357
Sargent, H., 360
Seashore, H., 181
Seward, G. H., 280
Shaw, F. J., 86
Sollenberger, R. T., 185
Stagner, R., 162
Strupp, H. H., 519
Triggs, F. O., 529
Wallace, S. R., Jr., 76
Weider, A., 89
Welch, L., 93
Werner, H., 165
Williamson, E. G., 362
Wolfe, J. B., 520
Worcester, D. A., 177
Youtz, R. P., 517

30n